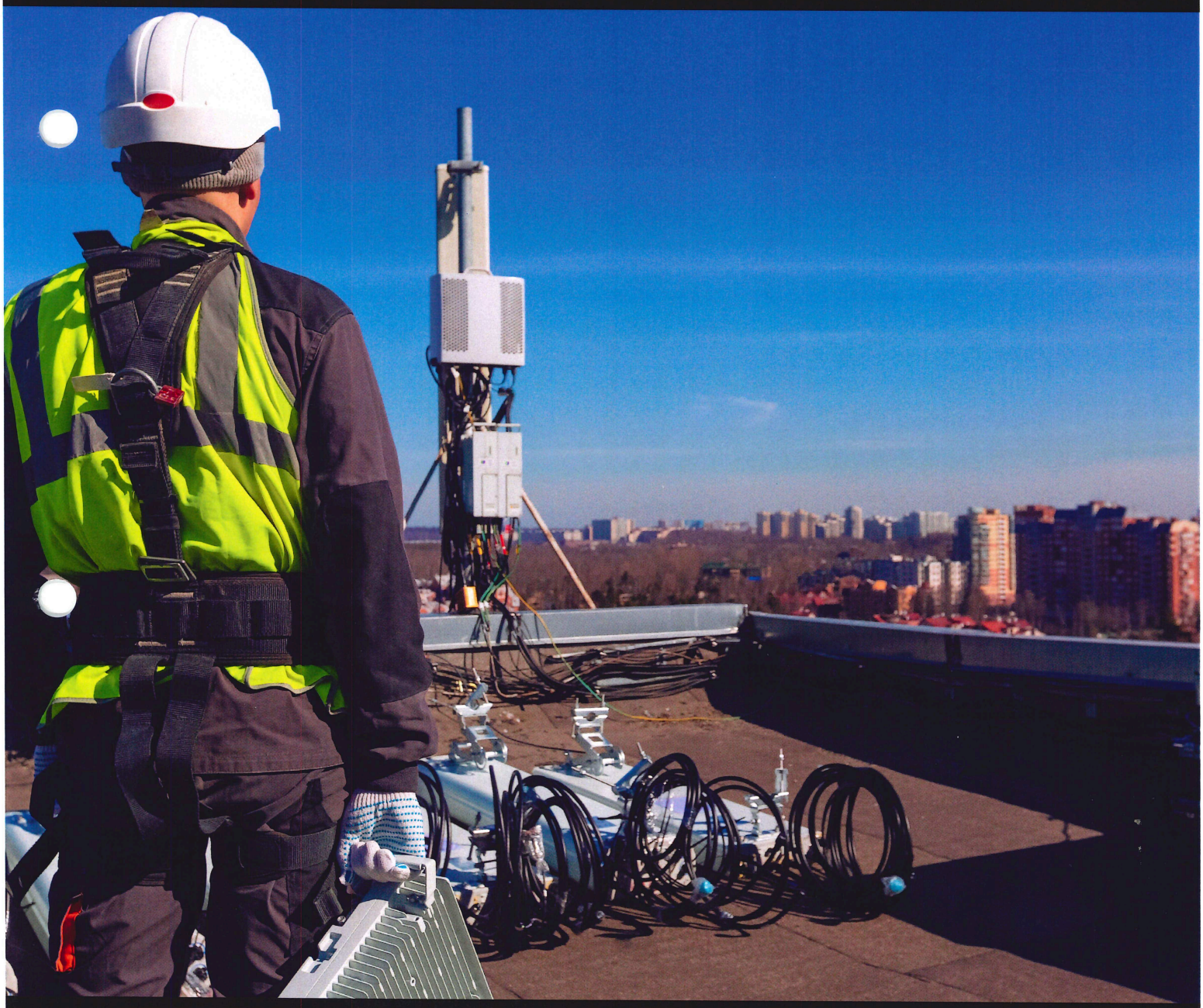
TAFE NSW

Solve fundamental electronic communications system problems



tafensw.edu.au

Learner Resource

Theory notes for Communications Fundamentals (UEEEC0063) (Version 1) by Barry Duncan (2023)

Copyright: TAFE NSW.

The original UEENEEH146A workbook (written by John Zervos and edited by Greg Denning) has been extensively modified and extended to create this workbook.

Table of contents

Lesson organiser 4

Required reading 5

Sections

1. An introduction to electronics communications systems 17
2. Electromagnetic wave propagation 31
3. An introduction to antennas (including the Yagi antenna) 41
4. Amplitude modulation and demodulation 63
5. Frequency modulation 93
6. FM demodulation 107
7. Introduction to receivers 123
8. The superheterodyne receiver 129
9. Optical communications principles 145

Ansers to review questions 165

Textbooks, eBooks and websites 183

Lesson organiser

The following is a guide only and will likely change to meet student needs.

Lesson	Section	Activity
Before starting (ideally)	Required reading	Work health and safety Using sustainable energy practices Communication and documentation Determining live testing/measuring requirements and isolating ccts
1	1	Exam Review (previous Unit) An introduction to electronics communications systems
2	2	Skill practise – An intro to the Biskit Trainer (Biskit Vol 1, Expt 2) Electromagnetic wave propagation
3	3	An introduction to antennas (including the Yagi antenna) Skill practise – Using the Biskit Trainer to model equations (Biskit Vol 1, Expt 3)
4 & 5	4	Amplitude modulation and demodulation Skill practise – Amplitude modulation (Biskit Vol 1, Expt 4) Skill practise – Amplitude shift keying (Biskit Vol 1, Expt 15 Part A) Skill practise – AM demodulation (Biskit Vol 1, Expt 6)
6	5	**Assessment – Refer to UAG for details** Frequency modulation
7 & 8	5 & 6	Exam Review Frequency modulation & demodulation Skill practise – Frequency modulation (Biskit Vol 1, Expt 9) Skill practise – Frequency shift keying (Biskit Vol 1, Expt 15) Skill practise - FM demod using the PLL (Biskit Vol 3, Expt 4)
9	7 & 8	Introduction to receivers The superheterodyne receiver Skill practise – Principles of superheterodyne (Biskit Vol 2, Expt 19)
10	9	Optical communications principles Revision
11 & 12	**Assessment – Refer to UAG for details**	

Required reading

These notes are provided to support your revision on aspects of this Unit that have been taught before in other Units (typically more than once) and will not be taught again but may be assessed.

This issue is explained in more detail on the following pages.

Introduction

As you probably know, units of competence that make up qualifications in the Electrotechnology Training Package (like CIII Electronics and Communications) prescribe what must be taught and assessed.

This is done using elements of competence together with performance criteria for each element. More detail on these is provided in the assessment requirements for the unit using two lists: one is the list of performance evidence; and the other is the list of knowledge evidence.

There are some performance criteria (together with performance evidence and/or knowledge evidence in the assessment requirements) that appear in many units and these include items on the following topics:

- work health and safety (WHS) using sustainable energy practices
- the appropriate use and completion of documentation
- appropriate communication with supervisors, customers and other stakeholders.

It makes sense that these performance criteria are common to multiple units because, taking WHS as an example, it's not reasonable to claim competence when the work was completed successfully but in an unsafe way.

Importantly, where all these topics have been comprehensively taught before, it is usually unnecessary to teach them a second or third time. That said, we are still expected to ask you to provide knowledge evidence and/or performance evidence possibly by answering questions on these topics in the theory test and/or demonstrating your skills on these topics during the practical tests.

For this reason, the following information and/or notes are provided for you to read and use to ensure that you can satisfy the assessment requirements for this unit.

Work health and safety (WHS)

When you undertake the in-class simulated workplace tasks and/or the practical test for this unit (EC0063), you're expected to demonstrate your knowledge of, and demonstrate your ability to carry out, the following performance criteria related to WHS and their associated knowledge evidence and performance evidence:

1.1 - Work health and safety (WHS) requirements and workplace procedures for a given work area are identified and applied (UEECD0007 references: 1.2, 1.3 & 1.4)

1.2 - WHS risk control measures and workplace procedures are followed in preparation for electronics communications system work (UEECD0007 references: 1.2, 1.3 & 1.4)

1.2 - Safety hazards not previously identified are risk assessed, noted on job safety assessments and established risk control measures implement (UEECD0007 references: 1.2, 1.3 & 1.4)

2.1 - WHS risk control measures and workplace for carrying out electronic work are followed (UEECD0007 references: 1.1,1.6, 2.1 and 2.4)

2.2 - Need to test and measure live work is determined in accordance with WHS requirements and workplace safety procedures (UEECD0007 references: 1.2, 1.3 & 1.4)

2.3 - Circuits/machines/plant are checked and isolated in accordance with WHS requirements and workplace procedures (UEECD0007 references: 1.2, 1.3 & 1.4)

3.1 - WHS risk control and work completion measures and workplace procedures are followed (UEECD0007 references: 3.1 to 3.4)

3.2 - Worksite is cleaned and made safe in accordance with workplace procedures (UEECD0007 references: 2.6)

Source: https://training.gov.au/Training/Details/UEEEC0063 As you may already be aware: WHS is taught and assessed in a dedicated unit, CD0007 (formerly E101A)

- CD0007 or E101A is an indirect pre-requisite unit for undertaking EC0063 (that is, you must have successfully completed CD0007 or E101A to be able to undertake EC0066 which is a pre-requisite of EC0063)

- The performance criteria listed above are underpinned by, and/or are identical in substance to, the performance criteria in CD0007 and E101A indicated in brackets above

- You have successfully met the assessment requirements for at least one or more of the following units which also include the performance criteria listed above: EC0066/H113A, EC0067/H139A, EC0069/H112A and EC0074/H114A.

Importantly, the knowledge and skills needed to perform these performance criteria are identical to those needed to repair resonance, amplifier and digital circuits operated from mains. This means that you already have the ability to demonstrate your knowledge of, and demonstrate your ability to carry out, these performance criteria in the context of this unit.

However, if you're unsure and would like to read more information about these performance criteria, please refer to your learner resources for CD0007/E101A that specifically address them.

Note: If you have not successfully completed CD0007 or E101A, please bring this to the urgent attention of the teacher before undertaking any of the simulated workplace tasks.

Using sustainable energy practices

When you undertake the in-class simulated workplace tasks and/or the practical test for this unit (EC0063), you're expected to demonstrate your knowledge of, and demonstrate your ability to carry out, the following performance criterion:

2.7 - Electronic problems are resolved without unnecessary damage to apparatus, circuits, the surrounding environment, or services using sustainable energy practices.

This includes knowing about sustainable energy principles and practices.

Sustainability is a much used word now that describes growing global response to the short and long term negative effects on the planet of certain human activities, particularly in relation to climate change and pollution. It's accepted that, without drastic action, both problems will likely make living on this planet more difficult (due to extreme weather and toxic environments) that, in turn, risk causing drought, famine, war and the failure of democratic governments and even whole nation states.

R J Fuller (Deakin University, 2005) describes four principles that underpin sustainability:

- Futurity - The concern for future generations
- Environment - The concern to protect the integrity of eco-systems
- Equity - The concern for the poor and disadvantaged
- Participation - The notion that individuals can participate in decisions affecting them

An essential element of the global response to climate change is the switch to sustainable energy. Energy is said to be sustainable if it "meets the needs of the present without compromising the ability of future generations to meet their own needs." Underpinning this are the sustainable energy principles which include (but are not limited to):

- Reducing our use of non-renewable energy sources (like the fossils fuels which includes coal, gas and oil)
- Increasing our use of renewable sustainable energy sources (like solar just to name one of many)

Fuller's fourth principle is important because it means that individuals, including you, can contribute directly and indirectly to sustainable energy principles by using sustainable energy practices where possible. Direct action that can be taken includes (but is not limited to):

- Switching off equipment not in use
- Turning off lights when leaving a room with nobody in it
- Opening windows to cool rooms instead of turning on room air-conditioning
- Using natural lighting
- Minimising waste
- Recycling waste
- Using public transport to and from work
- Drinking tap water

Indirect action that can be taken, includes (but is not limited to) encouraging your employer to:

- Use more energy efficient lighting Using more energy efficient equipment
- Reduce (in winter) and increase (in summer) the temperature setting of air conditioning so that to uses less energy
- Switch to using sustainably sourced and/or sustainably manufactured materials Switch to an SEU energy provider

With all that said, every workplace is different so what works with one may not work with another.

Communication and documentation

When you undertake the in-class simulated workplace tasks and/or the practical test for this unit (EC0063), you're expected to demonstrate your knowledge of, and demonstrate your ability to carry out, the following performance criteria related to workplace communication and documentation and their associated knowledge evidence and performance evidence:

1.4 - Nature of communication system problem is obtained from documentation or from work supervisor to determine scope of work to be undertaken

1.5 - Advice is sought from work supervisor to ensure the work is coordinated effectively with others

1.6 - Tools, testing devices and materials need to carry out the electronic work are obtained and checked for correct operation and safety.

2-6 - Unplanned situations are dealt with safely and with approval of authorised person/s in accordance with workplace procedures in a manner that minimises risk to personnel and equipment.

3.3 - Reports are written outlining electronic system problem and justifying solutions used

The specific knowledge required to satisfy these performance criteria is specific to each workplace setting. However, the skills are generic and so are transferable between workplace settings including simulated workplace settings at TAFE NSW. This means that, when you undertake the in-class simulated workplace tasks and the skills assessment(s), you are expected to demonstrate your ability in this regard.

Importantly, while undertaking the in-class simulated workplace tasks and/or the practical test for this unit at TAFE NSW, the following applies to key terms in the performance criteria above:

- "Authorised person" refers to the class teacher
- "Others" refers to the class teacher and other students in the class
- "Materials" refers to components and training aids needed to carry out the simulated workplace tasks as described in the skill practise exercises' equipment list
- "Workplace procedures" refer to the written and verbal procedures and instructions provided by the teacher for the safe and proper conduct of skill practise exercises
- "Tools, equipment and testing devices" refers to tools and test equipment needed to carry out the simulated workplace tasks as described in the skill practise exercises' equipment list and/or mounted on the workbenches

- "Documented... in accordance with workplace procedures" refers to the requirement for students to record their measurements, other results, and answers to questions on the skill practice exercises and/or any handouts that must be submitted to the teacher.

The following is the procedural requirement for undertaking the simulated workplace tasks that form the basis of skill practise exercises.

- Students must have permission from the teacher to start every skill practice exercise.
- Students must work individually on every skill practise exercise, or in pairs with the permission of the teacher (but never more than two per group).
- Students must perform a WHS risk assessment at the start of every skill practice exercise to ensure the safety of themselves and others.
- Students are to bring to the attention of the teacher any hazards that they identify for which there are they cannot implement appropriate control measures.
- Students are to collect equipment from the trolley in a safe and orderly manner having respect for the reasonable needs of other students.
- Students are to follow the written instructions in every skill practice exercises in correct order and in total including any qualifying notes.
- Students are expected to record their measurements, other results and answers to questions for every skill practice exercise and/or any handouts that must be submitted to the teacher.
- Students are expected to call the teacher to check their work when instructions direct them to do so.
- Students are expected to retake any measurements and reattempt and/or correct any of their answers to questions when reasonably directed to do so by the teacher.
- Students are not to return their equipment to the trolley or leave until they have completed the skill practice exercise and their work has been checked by the teacher.
- Students are expected to submit their work to the teacher by the due date when required to do so.
- At the completion of every skill practice exercise, students are expected to return all equipment to the trolley tidily (unless directed otherwise by the teacher) and ensure that the workspace is left cleaned and safe.

- Students are to bring to the attention of the teacher any faulty test equipment, tools and components.

Determining live testing/measurement requirements and isolating circuits

When using a multimeter or oscilloscope to test and measure electronics equipment, there are two important safety measures that can minimise the risk of your being fatally electrocuted.

The first involves ensuring that the equipment that you're testing or repairing (like a radio receiver) is plugged into a power point that is protected at the switchboard by a residual current device (RCD). You can check whether the building's powerpoints are RCD protected by looking inside the switchboard. Figure 1 below shows an example of what to look for.

Figure 1

Up close, RCDs look like the example in Figure 2 below and there should be at least one row of these devices in the switchboard mounted side-by-side. RCDs are easily recognised by the fact that they're switched devices and often they have an in-built test feature too.

Figure 2

RCDs have been designed expressly for the purpose of preventing people and animals from a sustained direct connection to mains which can be fatal.

If the switchboard is older it may use semi-enclosed cartridge fuses instead of RCDs and an example is shown in Figure 3 below. Importantly, being fuses, these devices are designed to protect equipment and prevent fires but they do not prevent people/animals from potentially fatal electrocution. You should avoid powering the equipment that you're testing/repairing from power points that are only protected at the switchboard by these fuses.

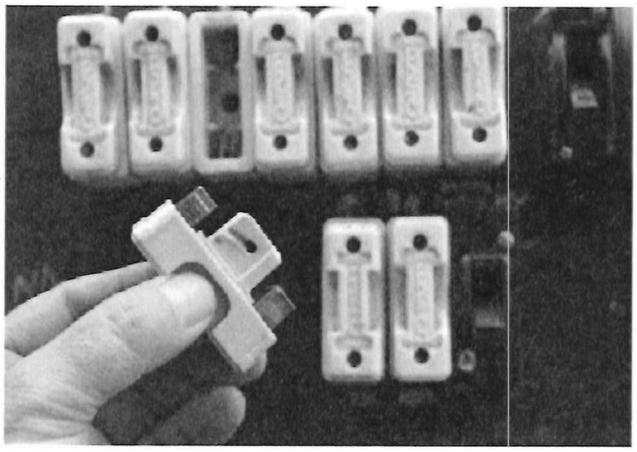


Figure 3

If you can't access the switchboard to see what's inside, or you know for certain that the power point is not RCD protected, then plug the equipment that you're testing into the powerpoint of a powerboard that is RCD protected instead. An example is shown in Figure 4 below - Notice the RCD at the left end (which includes a button for the test feature).

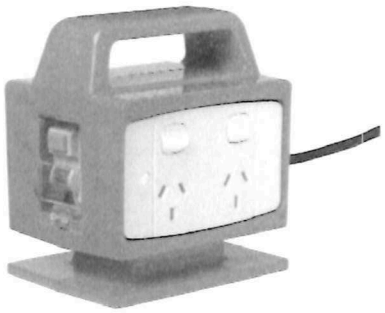


Figure 4

Take care when purchasing protected powerboards because there are many that offer "surge" and/or "overload" protection. While these features are useful, they don't provide protection from electrocution. Powerboards that are RCD protected will state this clearly and are usually a fair bit more expensive than simple surge and overload protected powerboards.

The second safety measure you must employ when connecting multimeters to live equipment, involves never doing so with your fingers touching any of the probe's metal parts. Figure 5 below shows the correct way to hold the probes. Notice that the person holds the probes behind their slip barriers.

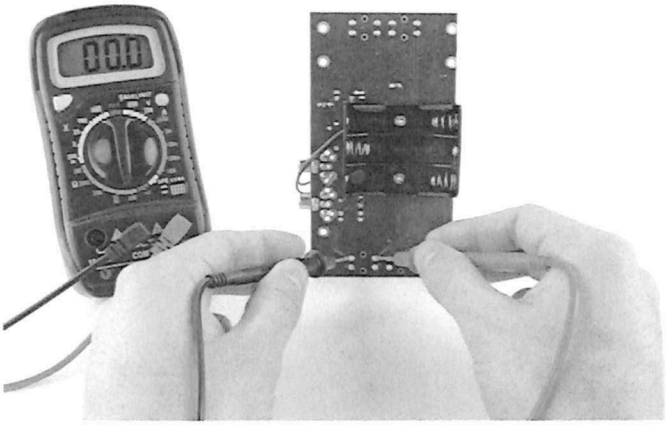


Figure 5

Student notes

Section 1 An introduction to electronics communications systems

This section aims to develop your general understanding of basic communication systems, the need for modulation and your knowledge of the radio frequency spectrum.

At the end of this section you should be able to:

a) State purpose of electronics communications systems.

b) Draw the diagram of a basic communications system and explain the purpose of each block.

c) Define the terms *simplex, half-duplex* and *full-duplex.*

d) With regard to the term *message*:

- Explain its meaning
- List four types of message in electronics communications
- State two other terms commonly used for "message"
- Draw a single tone and complex message in the time and frequency domains.

e) Give three reasons for using modulation in communication systems.

f) State the two types of modulation used in analog broadcast radio communication systems.

g) With regard to the term *noise*:

- Explain its meaning
- State point in the diagram of a basic communications system that is most susceptible to noise entering the system
- Explain *signal-to-noise ratio*
- Calculate the *signal-to-noise ratio* (in dB) given noise and signal powers/voltages.

h) List the bands in the radio frequency spectrum from VLF to EHF giving their full names.

i) State the frequency bands allocated to AM broadcast radio, FM broadcast radio and digital television (terrestrial).

j) Explain the term *channel bandwidth* and state the channel bandwidths for AM broadcast radio, FM broadcast radio and digital television.

Introduction

The purpose of any communication system is to convey information from a source to a destination. For example, in human conversation, a person (the source) speaks to pass on information to a listener (the destination). For another example, a newspaper is written by journalists (the source) to pass on the details of noteworthy events to readers (the destination).

Conveying information from a source to a destination is also the purpose of an electronic communications system. The main difference though is that, instead of using sound waves or paper, the information is sent using either electrical signals, electromagnetic waves or light (which is actually an electromagnetic wave too but at extremely high frequencies).

Electronic communications is integrated into all areas of the electronics industry, including the computing and electronic security industries. So, even if you don't choose to specialise in the area of electronic communications, a basic understanding of electronic communications principles is important to your role as an electronics technician.

The basic block diagram of an electronic communications system

All electronic communications systems can be represented by a common block diagram in Figure 1.

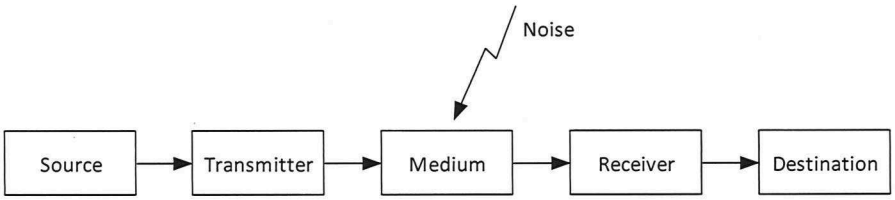


Figure 1 *The basic block diagram of an electronic communications system*

The **source** is the point of origin of the information that is to be sent over the communications system and is typically one or a combination of sound (speech, music, etc), text, images (drawings, photographs, etc), video, or measurement data. In electronics communications theory, the name for the signal containing the information is either, "intelligence signal", "baseband signal", or "message signal".

The **transmitter** has the necessary circuitry to facilitate the message signal being successfully propagated (transmitted) over the medium. Normally, this involves superimposing the message signal onto a higher frequency signal (usually a sinewave) called the "carrier" through a process called *modulation*. Modulation methods that you may already be familiar with include amplitude modulation (AM) and frequency modulation (FM).

The **medium** is the environment over which the signal is transmitted. This is normally either:

a) Copper or some other metallic conductor (for wire cables). For example, Ethernet cable for local computer networks.

b) Glass or clear plastic (for fibre optic cables). For example, the NBN telephone system (glass) and car networks (plastic).

c) The atmosphere and/or space (often called "free space"). For example, broadcast radio & tv, mobile telephone, Bluetooth, etc.

The **receiver** accepts the transmitted signal and has the necessary circuitry needed to create a copy of the original message that was used to modulate the carrier at the transmitter. This is called *demodulation*.

The **destination** is the where the recovered message signal ends up. This could be a loudspeaker, video screen, computer, or the audio/video/data inputs of some other system component.

Channels

The medium can be shared to allow the transmission of more than one information signal at a time. This can be done using either; time division multiplexing, frequency division multiplexing, or both. Technically speaking, users of the medium transmit over a "channel" which the ITU-T (an international telephony standards body) defines as "a means of transmission of signals in one direction between two points".

Transmission modes

The sophistication of a communications system can be described in terms of their transmission mode. The three transmission modes are:

Simplex – An electronic communications system that allows for information to be sent between two points in one direction only (eg broadcast radio and television).

Half-duplex – A system that allows for information to be sent between two points in both directions, but not at the same time (eg amateur radio).

Full-duplex – A system that allows for information to be sent between two points in both directions simultaneously (eg mobile phones).

Message signals

Most message signals such as speech, music, video and data are complex signals and so consist of thousands and even millions of sinewaves at frequencies within a band. For example, speech sounds produce fundamental frequencies together with significant harmonics at frequencies from about 100Hz up to about 4kHz (depending on what you call "significant"). Music can produce harmonics up to the teens of kilohertz, though this very much depends on the instruments (eg guitar, drum, etc) so that figure is indicative. Data has significant harmonics up to five times the signal's clock rate (this is Eric Bogatin's Rule of Thumb #2).

For ease of explanation when talking about communications systems, and for ease of student understanding, it is common practice to use a single simple sinewave for the message and that is done throughout this book.

Another common practice when talking about communication systems involves representing signals in both the time domain and frequency domain. You may recall learning about these two representations before.

Figure 2 below shows a 10Vp 1kHz sinewave shown in the time domain. The vertical axis is used to show the signal's amplitude and the horizontal axis represents time. This shows how the sinewave's voltage changes with time and is the same as what you would see on the display of an oscilloscope.

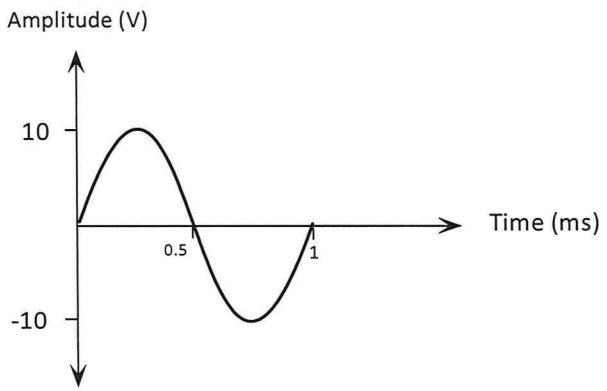


Figure 2 *A 10Vp 1kHz sine wave shown in the time domain*

Figure 3 below shows the same 10Vp 1kHz sinewave shown in the frequency domain. Notice that the vertical axis is still used to show the signal's amplitude but the horizontal axis represents frequency (instead of time). This shows the signal's position in the spectrum of signal frequencies and is the same as what you would see on a spectrum analyser.

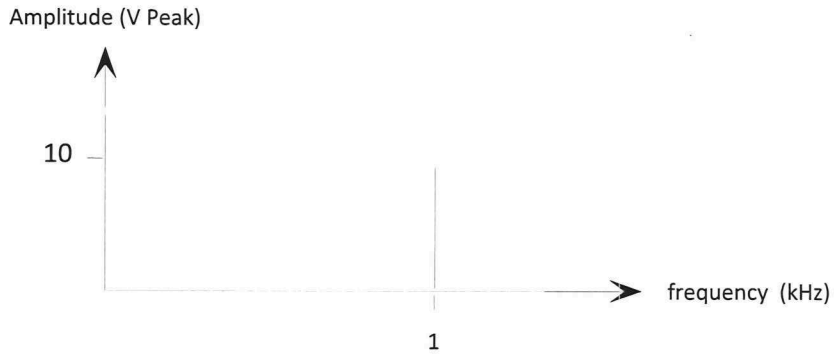


Figure 3 *A 10Vp 1kHz sine wave shown in the frequency domain*

(Strictly speaking, Figure 3 shows what a 10Vp 1kHz sinewave would look like on the display of an ideal spectrum analyser. The implementation of spectrum analysis is challenging and so the sinewave would be represented on the spectrum analyser's display as more of a narrow peak with very, very steep sides rather than a straight line.)

As already mentioned, complex message signals consist of hundreds, thousands and even millions of sinewaves between a minimum and maximum frequency. Represented on the frequency domain, these signals would look like a block and an example is shown in Figure 4 below.

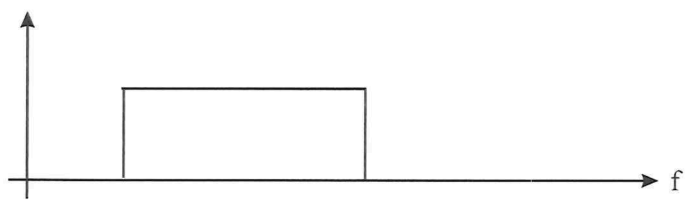


Figure 4 *The frequency domain representation of a complex message*

Note: This drawing is representative of the message consisting of many sinewaves at different frequencies but doesn't necessarily provide an indication of the magnitudes of those sinewaves relative to each other.

The need for modulation

Message signals are not suited for direct radio transmission. That is, it doesn't work very well if you simply power amplify the message signal and connect it to antenna. This is because message signals are relatively low frequency and antennas have poor radiation efficiency at lower frequencies. Moreover, as you will see, the antenna would also be required to be hundreds of meters, even kilometres, long.

There is further problem. Directly transmitting the message signal would mean that there could only be one operator transmitting any type of signal in the local area that the signal reaches (that is, within the transmission's radio horizon). If another baseband signal was directly transmitted, it would mix with the first and the two would be irretrievably mixed at the receiver. You can't unscramble an egg, as they say.

For these reasons, the message signal is normally modulated (that is, encoded or superimposed) onto a higher frequency carrier signal (most often a sinewave). The higher frequency carrier allows the antenna to be more efficient, shortens the antenna length and permits multiple users in the one area (a concept called frequency division multiplexing).

When you tune into a radio or TV station, you're actually tuning in to its modulated carrier frequency which the receiver demodulates to recover the original message.

Types of modulation method

There are many types of modulation method used in electronic communications. Some of the more common ones are briefly described below, with fuller treatment of AM and FM given throughout the remainder of this workbook.

Amplitude Modulation (AM) – The amplitude of a high frequency sine wave carrier is modulated (varied) in accordance with the instantaneous amplitude of the message signal.

> "Full AM" is referred to as Double Sideband Full Carrier (DSBFC) AM. Variations of this are Double Sideband Suppressed Carrier (DSBSC) and Single Sideband Suppressed Carrier (SSBSC). More in another chapter.

Frequency Modulation (FM) – The frequency of a high frequency sine wave carrier is modulated (varied) in accordance with the instantaneous amplitude of the message signal.

Phase Modulation (PM) – The phase of a high frequency sine wave carrier is modulated (varied) in accordance with the instantaneous amplitude of the message signal. Phase modulation is similar in many ways to frequency modulation.

Quadrature Amplitude Modulation (QAM) – This is a combination of phase and amplitude modulation.

Noise

According to the ITU-T, *noise* is an electromagnetic phenomenon, having components in the radio frequency range, which may or may not be carrying information, that is superimposed or combined on the wanted signal and can interfere with the quality of the recovered message at the destination.

You're probably most familiar with noise as buzzing, crackling, pops or hiss out of speakers. But noise can corrupt data causing problems with digital systems such as broadcast tv.

Interestingly, this definition of noise includes otherwise wanted signals. To explain, if you're listening to one radio station and you're too near the transmission antenna of another radio station, you may hear the second radio competing with (or even replacing!) the one you want to listen to. Although the second radio station is the wanted signal for other listeners, it's noise to you.

While noise can enter the system at any point, the medium is the most vulnerable because the signal is usually at its smallest. Moreover, for wireless communications, free space is a very noisy environment, electromagnetically. So the medium is usually the greatest source of noise in for most communications systems.

A relative measure of the amount of noise in a system is given by its *signal-to-noise ratio* (SNR), which is the ratio of desired signal power to noise power expressed in decibels. That is:

$$SNR = 10 \times Log\left(\frac{P_{Signal}}{P_{Noise}}\right)$$

Where: SNR = Signal-to-noise ratio (dB)

P_{Signal} = Signal power (W)

P_{Noise} = Noise Power (W)

Calculated using measured voltages:

$$SNR = 20 \times Log\left(\frac{V_{Signal}}{V_{Noise}}\right)$$

Where: V_{Signal} = Signal voltage (V)

V_{Noise} = Noise voltage (V)

The radio frequency spectrum

The range of frequencies used for radio communications is referred to as the *radio frequency spectrum*. The entire radio frequency spectrum is broken up into groups for ease of identification and these are listed in Table 1.

Table 1 *Division of the radio frequency spectrum*

Abbreviation	Name	Bandwidth	
VLF	Very Low Frequency	3kHz to 30kHz	
LF	Low Frequency	30kHz to 300kHz	
MF	Medium Frequency	300kHz to 3000kHz	→ KNOW THIS
HF	High Frequency	3MHz to 30MHz	
VHF	Very High Frequency	30MHz to 300MHz	→ KNOW THIS
UHF	Ultra High Frequency	300MHz to 3000MHz	→ KNOW THIS
SHF	Super High Frequency	3GHz to 30GHz	
EHF	Extra High Frequency	30GHz to 300GHz	

→ MICROWAVES

Different parts of this spectrum are used for different applications and radio transmission technologies.

For wireless communications, each channel is assigned a carrier frequency by the Australian Communications and Media Authority (ACMA). The assigned channel frequency sits inside a section (band) of the radio spectrum that has been set aside for that particular type of communication or broadcast service.

Radio and television broadcast frequencies

AM broadcast radio uses a portion of the MF band on channel frequencies between 531kHz and 1602kHz inclusive (with extensions up to 2000kHz for community/narrowcast radio services). For example, Sydney AM radio station 2GB transmits on 873kHz. (Using rounding, it's often said that the channel frequencies are from 530kHz to 1600kHz and this would be an acceptable answer in an assessment.)

AM broadcast frequencies are allocated by the Australian Communications & Media Authority (ACMA) to approved licence holders in 9kHz intervals across the band and never any frequencies in between. So there's an AM radio station transmitting on 531kHz and another one transmitting on 540kHz but no Australian radio stations are transmitting on 532kHz, or 533kHz, or 534kHz and so on. (This is done to avoid inadvertent interference of one radio station by another should its signal travel extraordinarily long distances, which can happen under certain conditions.)

FM radio stations use the VHF band with assigned station frequencies from 88.1MHz to 107.9MHz inclusive. For example, Sydney station Nova 969 is on 96.9MHz. Again, using rounding, it's often said that the channel frequencies are from 88MHz to 108MHz. FM broadcast frequencies are allocated to approved licence holders in 200kHz intervals across the band.

You can check the full list of Australian radio stations here:

https://en.wikipedia.org/wiki/List_of_radio_stations_in_Australia

DAB+ radio – Not tested
DAB stands for digital audio broadcasting and the plus symbol is a reference to an improved version of the earlier DAB standard. In Australia, the band of frequencies allocated to DAB+ is 174MHz to 230MHz which overlaps with the terrestrial TV broadcast band. See below.

Digital television (DVB-T) is broadcast in the VHF band (channel frequencies from 177.5MHz to 226.5MHz) and at the bottom end of the UHF band (channel frequencies from 529.5MHz to 697.5MHz). For example, Channel 10 transmits on 212.5MHz. TV broadcast frequencies are allocated to approved licence holders in 7MHz intervals across the band.

You can check the full list of Australian VHF & UHF TV channel frequencies here:

https://en.wikipedia.org/wiki/Television_channel_frequencies

Interesting fact – Not tested

When analog television system was finally switched off, digital tv stations at the top end of the tv band were "restacked" onto lower frequency channels to free-up spectrum. (So channels 53 to 69 where effectively shut down. On this point though, you shouldn't get confused between these channel numbers and the LCN numbers that you type on your remote control to select digital tv channels.)

This restacking was called the "digital dividend" as it allowed the federal government to auction the newly freed-up UHF spectrum up to 820MHz. Most successful bidders used the chunk of spectrum they bought for high-speed 4G broadband services. What does that really mean? Several 4G channels are located at the upper end of the old tv band from 700MHz. Mobile phone providers using these channel frequencies bought some of the newly released spectrum to increase the bandwidth of their channels which allowed them to increase data rates.

Does this mean there'll be a digital dividend in another part of the spectrum now that 3G has been turned off? No. At least one mobile phone provider turned their 3G channels into 4G channels so those channel frequencies are still being used.

Will there be a digital dividend when 5G is cheap enough and overtakes 4G as the preferred technology by the overwhelming majority of customers? I don't know. Your author has done a quick bit of research online and there doesn't appear to be any talk of a digital dividend from 5G at the time of writing. If it were to happen, it would probably be at least a decade away.

Channel bandwidth

As previously mentioned, the stated frequency of a TV or radio channel is actually the frequency of the channel's carrier. However, this is not the only sinewave that is transmitted. When carriers are modulated, a range of new sinewaves are generated that straddle the carrier and these are called the *sidebands*.

For all modulation schemes, some or all of the intelligence in the message resides in the sidebands. Consequently, every radio and TV channel must be allocated a band of frequencies within which to transmit their modulated signal. This is referred to as channel bandwidth.

AM radio stations are allocated a channel bandwidth of 17kHz (ie ±8.5kHz either side of the carrier). FM radio stations are allocated a channel bandwidth of 200kHz (though the actual bandwidth of broadcast FM radio transmissions is not meant to exceed 180kHz so this figure includes "guard bands" at each end of the modulated signal.) TV stations are allocated a bandwidth of 7MHz.

A decision about the bandwidth of channels in any band is a compromise. Wider channel bandwidths allow more information to be sent over each channel and this is a good thing because this allows for higher quality (or higher "fidelity") sound to be transmitted, or higher bit rates where the message is data. But wider channel bandwidths reduce the total number of channels (ie users) that can operate in the band. Conversely, smaller bandwidths allow less information to be sent (so, lower quality sound or lower data bit rates) but it allows for more channels on the band.

An example of this compromise – Not tested

Voice telephony is an interesting example of this compromise. A decision was made in the early days of telephony to narrow the channel bandwidth to 3.1kHz (ie the range was from 300Hz to 3.4kHz). This made perfect sense at the time because it maximised the number of channels (ie telephone conversations) on the line while transmitting only the most significant sinewaves in speech to allow the caller to be easily recognised and understood by the listener.

But, you've probably noticed that, when you're put on hold and the hold music is from a local FM radio station, it sounds terrible over the phone compared to listening to the same FM station on a radio itself. That's because the channel bandwidth in telephony doesn't permit the transmission of enough of the sinewaves in the music signal to make it sound like it does live or from an FM radio station (which can transmit sinewaves in message signals up to 15kHz).

Review questions

Answer these questions to check your understanding of what you have learnt from this chapter. Doing this will help you prepare for the knowledge assessment.

1. Sketch the basic block diagram of a communications system and correctly label each block.
2. Give an example for each block in the basic block diagram of a communications system.
3. What are the terms "baseband signal", "message signal" and "intelligence signal" used to describe?
4. What is the definition of "noise" in an electronic communications system?
5. Why is the medium in the block diagram of a communications system the stage that is most susceptible to noise?
6. Describe the term "modulation" in the context of electronic communications.
7. Give two reasons for using modulation (instead of transmitting baseband signals).
8. Write the long names for each of the following abbreviations:
    - VLF
    - DSBFC
    - UHF
    - SSBSC
    - MF
    - FM
    - VHF
    - PM
    - HF
    - AM
9. State the channel bandwidth for AM and FM broadcast radio and television in Australia.
10. Calculate the SNR if the signal power is 150mW and the noise power is 800µW.
11. What's the important difference between the equations when calculating SNR in decibels using power figures versus using voltage figures?
12. Sketch a 10V peak, 5kHz sinewave in both the time domain and the frequency domain.
13. Sketch a complex baseband signal in the frequency domain that has its lowest frequency at 200Hz and highest frequency at 4kHz.

Student notes

Section 2 Electromagnetic wave propagation

This section aims to develop your understanding of basic electromagnetic wave propagation.

At the end of this section you should be able to:

a) Describe the basic properties and characteristics of an electromagnetic (EM) wave.

b) Calculate the wavelength of an EM wave.

c) State and describe the three modes of EM wave propagation from an antenna.

d) State an advantage and a limitation of each of the three modes EM wave propagation.

e) State the predominant mode(s) of propagation and communication application(s) at MF, HF, VHF and UHF frequencies.

f) Describe the following with regard to satellite communications:

- The meaning of a geostationary orbit
- The frequency bands used for uplink transmissions
- The common name used for signals at these frequency bands

Electromagnetic waves

You have learnt in earlier Units that a **magnetic field** is produced around any conductor carrying an electric current. You have also learnt that an **electric field** is produced between any two conductors or points in a circuit that have a potential difference between them. The energy in these fields is normally returned to the supply when the power is switched off and these fields collapse.

When the conditions are right, some of this electric and magnetic energy is not returned to the circuit and is instead radiated (emitted) as electromagnetic energy. When the source is AC, the electromagnetic energy is radiated in sinusoidal waves having the same frequency as the source.

As its name implies, an electromagnetic wave consists of two fields; an electric field (known as the E field) and a magnetic field (known as the H field). These fields are perpendicular (90°) to each other and both are perpendicular to their direction of travel. This is shown in Figure 1.

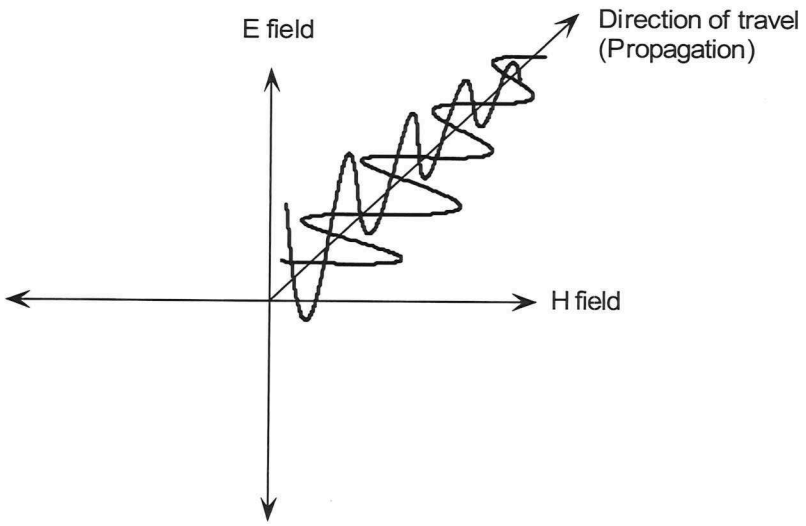


Figure 1 *Relationship between the electric and magnetic fields in an electromagnetic wave*

In electronic communications, electromagnetic waves are often called radio waves.

Electromagnetic waves are said to be polarised in the direction of their electric (E) field. If the E field is vertical, as in Figure 1, then the electromagnetic wave is said to be vertically polarised. If the E field is horizontal, the wave is said to be horizontally polarised.

Radiated electromagnetic energy is often undesirable as it can cause radio frequency interference (RFI) in other circuits. However, at the antenna of a radio transmitter, electromagnetic radiation is exactly what is needed.

Wavelength

Electromagnetic waves in a vacuum and free space travel at a constant velocity equal to the speed of light (in fact, light is an electromagnetic wave as well, just an extremely high frequency). The speed of light is 3×10^8 m/s (metres per second, also written as ms^{-1}). That's 300,000 km per second.

(To put that in perspective, the earth is 149.33 million kilometres away from the sun but its light takes just eight minutes to reach us.)

As electromagnetic waves are travelling, and as it takes a finite time for them to complete a cycle, one cycle of any electromagnetic wave must take a certain physical distance to complete.

If the frequency of the electromagnetic wave is known, its wavelength can be calculated using the following equation:

$$\boxed{\lambda = \frac{c}{f}}$$

Where: λ (pronounced "lambda") is the wavelength in meters (m)

c is the velocity (speed) of light which is 3×10^8m/s (in free space)

f is the frequency of the waveform in hertz (Hz)

For example, the wavelength of a 1MHz electromagnetic wave is:

$$\lambda = \frac{c}{f} = \frac{3\times10^8}{1\times10^6} = 300m$$

Types of electromagnetic wave propagation

There are three modes of EM wave propagation from the transmitting antenna together with a fourth option for satellite communications. These are:

- Ground waves
- Space waves (which, confusingly, don't travel into space!)
- Sky waves
- Satellite communications

Ground waves

Ground waves (also called surface waves) are EM waves that hug the earth's surface. In fact, the currents that help produce the ground waves actually travel in the earth's surface and this is the cause of the hugging effect.

To transmit ground waves, the antenna must be vertically polarised so that the E field is vertically polarised to prevent it from shorting out on the earth's surface (which can't happen to the H field).

Ground waves travel best along salt water. They also travel well along fresh water and over ground that is flat, with good agricultural soil and/or is marshy. Ground waves don't travel as readily along dry, sandy soil or along rough terrain.

The advantages of using ground waves for radio communications include:

- They are very reliable as they are not affected by changes in the ionosphere (though transmissions along water can be affected by changes in weather)
- They are capable of over the horizon communications

Ground losses increase rapidly with frequency, especially above 2MHz to 3MHz (depending on which book you read). So ground wave frequencies are best suited transmissions in the medium frequency (MF) band and lower.

AM broadcast radio transmits signals using ground waves. That's why their antennas are vertically polarised and the antenna tower that you see is the antenna itself (whereas, FM radio towers are not the antenna but only a mast for the antenna which is mounted near the top). To make best use of ground waves, AM broadcast radio uses Marconi antennas (more later).

Ground waves are used in some marine communications and are the only option for communications with submarines (other than satellite communications).

Space waves

Space waves don't hug the ground like ground waves do and instead travel above the ground. Space waves are the most widely used mode of antenna propagation because they are better suited to high frequency propagation. (Remember, higher frequencies are generally better to use because the amount of energy radiated from an antenna increases as the frequency increases. Also, as you will see, higher frequencies require smaller antennas for optimal transmission.)

Space waves can be either direct or reflected as shown in Figure 2.

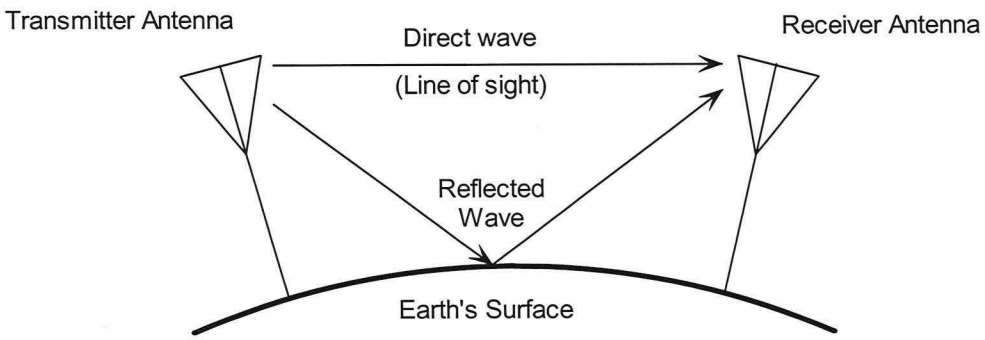


Figure 2 *Direct and reflected space waves*

Space waves travel in straight lines in multiple directions away from the transmitting antenna. This fact causes a practical limitation to their use. As the distance between the transmitter and receiver antennas increases, they move further around the curvature of the Earth. When they're far enough apart, all of the transmitter's space waves pass directly over the receiver's antenna.

For space wave communications, you should be able to actually "see" (possibly with the aid of a telescope) the transmitter antenna from the receiver antenna and vice-versa (assuming no hills or other objects are in the way). It's for this reason that space waves are said to be "line-of-sight" communication.

The obvious solution to the line-of-sight limitation due to the Earth's curvature is to increase the height of the transmitter's antenna, increase the height of the receiver's antenna, or both. Broadcast radio systems using space waves (FM radio and television) mount the transmitter's antenna on top of a tall tower to increase the reach of their signals. If you look around on your travels, you'll notice that tv antenna heights in Sydney's outer suburbs are often higher than for inner suburbs.

Space waves are generally used for VHF and UHF transmissions including broadcast FM radio and television. One of a variety of dipole antennas are for these applications including the Yagi antenna (more in a later chapter).

Generally speaking, radio waves pass through solid objects. But, they can be reflected by certain materials such as water and metal. So, space waves can reflect (or bounce) off the ground or other objects like buildings, water towers and even aeroplanes. Reflected space waves can cause reception problems because two versions of the same signal arrive at the receiver with the reflected wave being slightly delayed (or out of phase) due to having travelled further to get to the antenna.

This problem was seen on old analog television receivers as "ghosting". Digital television receivers are less affected by reflections because the reflected signal is often too weak to corrupt the bits sufficiently to confuse the receiver's digital electronics. Moreover, the data can be processed to eliminate problems caused by delayed signals.

Sky waves

While increasing the height of the transmit and/or receive antennas is a good solution to the line-of-site limitation of space wave propagation, it becomes impractical above certain heights and even a hazard to air traffic. This limits the transmission distance to about 30km over flat ground without the use of repeaters (the quoted figure is indicative).

Much further distances can be achieved by directing space waves towards the sky and so, by definition, they become "sky waves". This works because of a region in space approximately 50km to 400km above the ground called the *ionosphere*. Sky waves at the right frequency and directed at the right angle are **refracted** back towards earth at distances beyond line-of-sight.

Sky wave propagation is best suited to frequencies in the HF band (3MHz to 30MHz). If the frequencies of the transmitted sky waves are too low, they can be **reflected** back to the ground by a lower level of the ionosphere (the D layer) that exists during daylight hours reducing the propagation distance. If the frequencies of the transmitted radio wave are too high, they can pass through the ionosphere and continue out into space.

It's difficult to put a specific figure on the distances that can be achieved by sky wave propagation because the necessary refraction is a function of the frequency used and is affected by solar activity which varies throughout the day, between seasons, and with the sun's 11-year sunspot cycle.

Interestingly, when the refracted sky waves hit the ground they may be reflected (bounced) back towards the ionosphere to be refracted again in an action called *skipping*. This skipping action is illustrated in Figure 3 and can facilitate communications at very long distances.

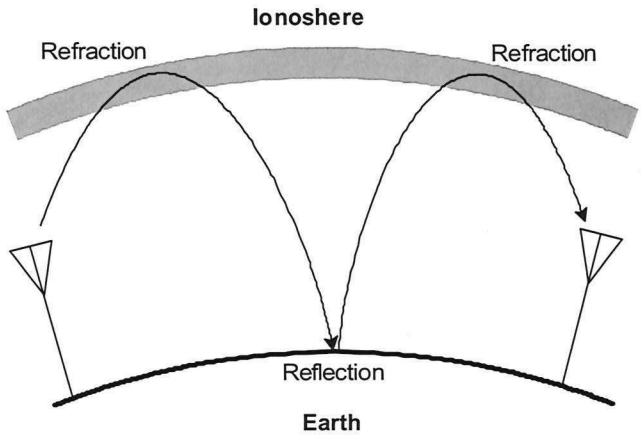


Figure 3 *Sky wave propagation involving skipping*

Interesting facts – Not tested

To explain why it's difficult to be exact about the distances that can be achieved using sky wave propagation, you need to know a little more about the ionosphere.

The ionosphere consists of three layers designated D, E and F from lowest in altitude to highest. The F layer is further subdivided into the F_1 (lower) and F_2 (upper) layers.

The D layer is closest to the ground and is at maximum density during the daylight hours of summer. It's less dense during the day in winter and effectively disappears at night all year round. The D layer can reflect VLF and LF frequencies and absorb MF and HF frequencies. So the D layer doesn't play a significant part in sky wave propagation. (That said, at night it's possible to hear AM radio stations at long distances due to the absence of this layer permitting inadvertent sky wave propagation of their MF signals. Your author has been driving around Sydney at night and has heard AM radio stations from Victoria and from somewhere near Tamworth!)

The E layer is the middle layer in the ionosphere and can refract higher frequencies up to about 20MHz (ie frequencies towards the top of the HF band). Like the D layer, the E layer effectively disappears at night.

The F layer is the highest layer and is present both day and night. During daylight hours, the sun's radiation separates the F layer into the F_1 and F_2 layers. Ionisation is at a maximum at noon. The F layer is responsible for long distance transmissions at frequencies up to 30MHz.

Satellite communications

The distance limitations of space wave and sky wave communications made satellite communications an obvious choice for long distance communications in the latter half of the 20th century. (That said, the relatively recent use of optical fibre technology for long distance telecommunications has seen a relative decline in use of satellites for telecommunications.)

Communications satellites are typically placed into geostationary orbit (ie the satellite orbits the earth above a fixed point on the ground) at an altitude of 35,900km. The ground based transmitting antenna sends a highly directional signal through the ionosphere to the satellite's receiving antenna. The satellite re-amplifies the received signal and transmits it (using a different carrier frequency) either: a) back to earth; or b) to another satellite.

These days, the frequencies used for the "uplink" communications range from around 1GHz to 40GHz, putting them in the UHF, SHF and EHF bands (which are generally referred to as microwaves because of the shortness of their wavelengths). These carrier frequencies pass through the ionosphere more easily than lower frequencies and they also permit wide bandwidths to allow the multiplexing of many different signals.

Satellite communications allow very long (transoceanic) distances to be achieved. Just three geostationary satellites can allow communications to provide global coverage. Satellites can also be used to broadcast signals to entire regions. For example, the Optus D1 satellite can provide television signals to the whole of Australia and New Zealand.

Review questions

Answer these questions to check your understanding of what you have learnt from this chapter. Doing this will help you prepare for the knowledge assessment.

1. What is an electromagnetic wave?
2. Electromagnetic waves are polarised in the direction of which field?
3. What happens to electric and magnetic fields that are not returned to the circuit when they collapse?
4. Calculate the wavelength of an electromagnetic wave having the following frequencies:
   a. 20MHz
   b. 500MHz
5. What are ground waves?
6. Which frequency bands are best suited to ground wave transmissions?
7. Give an example of a communications system that makes use of ground waves.
8. Space waves are said to be "line-of-sight". What does this mean?
9. Which frequency bands are best suited to space wave transmissions?
10. In what way can reflected space waves become a problem?
11. What is the ionosphere and why is it so important in wireless communications?
12. What are sky waves?
13. Explain why longer communications distances are possible with sky waves compared to space waves.
14. What is a communications application for sky waves?
15. Satellites used in electronic communications are geostationary. What does this mean?
16. Explain the general principle of satellite communications.
17. Explain why satellite communications can achieve the longest transmission distances.

Section 3 An introduction to antennas (including the Yagi antenna)

This section aims to develop your understanding of basic antenna design and operation principles.

At the end of this section you should be able to:

a) Describe the key physical features of the half-wave dipole and quarter-wave monopole antennas.

b) Determine the ideal length of the half-wave dipole and quarter-wave monopole antennas for a given carrier frequency.

c) State which type of electromagnetic wave propagation is primarily used by the half-wave dipole and the quarter-wave monopole antennas.

d) State the meaning of the following terms with regard to antennas: *polarisation*, *reciprocity*, *radiation pattern*, *directivity* and *gain*.

e) Sketch the typical radiation patterns for horizontally polarised half-wave dipole antenna and the vertically polarised quarter-wave monopole antenna with ground plane.

f) State two advantages of using directional antennas like the Yagi antenna.

g) Draw the basic construction of a three-element Yagi antenna

h) Calculate the key dimensions of a Yagi antenna optimized for a given carrier frequency.

i) State the hazards associated with working on and/or near antennas and describe the necessary control measures.

Introduction

Antennas are intended to either convert electrical energy to electromagnetic energy (at the transmitter) or vice versa (at the receiver) and so they can be thought of as a transducer.

Before discussing antenna design, there are a few important basic facts about antennas that you should know.

First, the amount of energy radiated from any antenna increases as the frequency of the transmission signal increases. (Recall that this is one of the reasons for modulating a higher frequency carrier with the message signal instead of directly transmitting the message signal.)

Second, the characteristics of an antenna are essentially the same whether sending or receiving electromagnetic energy. This means that, when an antenna is optimally designed to transmit certain signals in certain directions, it's also optimally designed to receive the same signals from the same directions. And the reverse is also true. This called antenna "reciprocity".

Third, dipole and monopole antennas (more on those next) can be mounted vertically or horizontally, giving rise to either a vertically or horizontally polarised radio wave respectively. Importantly, to maximise the size of the received signal, the receiver's antenna must be mounted in the same polarisation.

Another important fact – Not tested

The received signal strength of an antenna is described in terms of the electric (E) field's strength. This is calculated by dividing the received induced signal voltage by the antenna length. For example, if the antenna has an induced signal voltage of 20μV and is 2m long, the field's strength is 10μV/m.

Efficient antenna design

Generally speaking, a single piece of wire of any length will work as an antenna. However, it may not do so efficiently, either in terms of transmitting signals or receiving them. Efficient antenna design is about:

- Converting the maximum amount of electrical energy from the transmitter to EM radiation for propagation into free space; or
- Harvesting the maximum amount of the desired channel's EM signal in free space to convert it to an electrical signal for use by the receiver.

(Remember, antenna reciprocity means that achieving one achieves the other.)

Efficiency matters with antennas. When transmitting radio signals, the greater the power of a radiated EM wave, the greater the distance it will travel (all other factors being equal). This means that signals can travel further for the same power, or less power is needed to transmit signals a certain distance. When receiving signals, the more energy in the EM wave that an antenna can harvest, the further it's possible to reliably receive signals using the antenna.

The half-wave dipole

The half-wave dipole (also known as the Hertz antenna) is an efficient antenna design and is one of the most widely used types of antenna for transmitting signals at frequencies in the HF, VHF and UHF bands.

As its name implies, the half-wave dipole involves two conductors of equal length arranged axially (often also said to be "collinearly") having a total length that is equal to half the wavelength of the desired carrier frequency of the modulated signal being transmitted and/or received. This is shown in Figure 4 below. A large gap is shown between the conductors to emphasise their physical separation. However, in practice, the gap should be small.

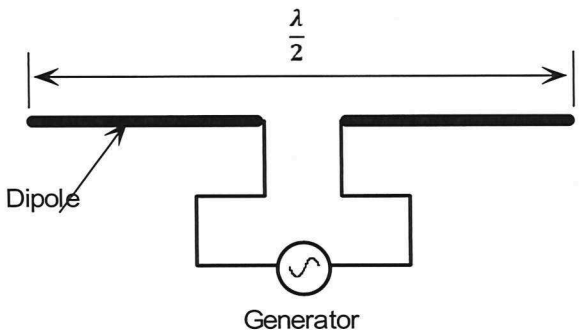


Figure 4 *The half-wave dipole or Hertz antenna*

To determine the overall length of the half-wave dipole, you simply determine the wavelength of the carrier signal you want to transmit and/or receive then divide that figure by two.

For example, the length of a half-wave dipole antenna tuned to Sydney TV Channel 10 on 212.5MHz is:

$$\lambda = \frac{c}{f}$$

$$\lambda = \frac{3 \times 10^8}{212.5 \times 10^6}$$

$$\lambda = 1.41m$$

Therefore, the half-wave dipole's length should be:

$$l = \frac{\lambda}{2}$$

$$l = \frac{1.41m}{2}$$

$$l = 0.705m$$

For another example, the length of a half-wave dipole antenna tuned to an AM radio station transmitting on 531kHz (which is the lowest allowable broadcast AM carrier frequency in Australia) is:

$$\lambda = \frac{3 \times 10^8}{531 \times 10^3} = 565m$$

$$l = \frac{565}{2} = 282.5m$$

This second example demonstrates a practical limitation of the half-wave dipole. As the frequency of an EM wave decreases, its wavelength increases. So, if the carrier frequency is low enough, the antenna becomes too long to be practically mounted either vertically or horizontally. (To put 282.5m from the last example into perspective, Centrepoint Tower is 305m high.)

Why the half-wave dipole is an efficient design – Not tested

The following is the concise explanation for why the dipole makes for an efficient antenna. An AC electrical signal on a conductor has a physical length as well as a period. Unsurprisingly, that length is the same as the wavelength of an EM wave from an antenna at the same frequency.

As the two ends of a half-wave dipole are open-circuit, the current is zero at the ends and the voltage is maximum. If the half-wave dipole's length is equal to half the carrier's wavelength then, in the middle of the dipole - which is a quarter of a wavelength back from both ends - the current must be maximum and the voltage zero. This is shown in Figure 5 below.

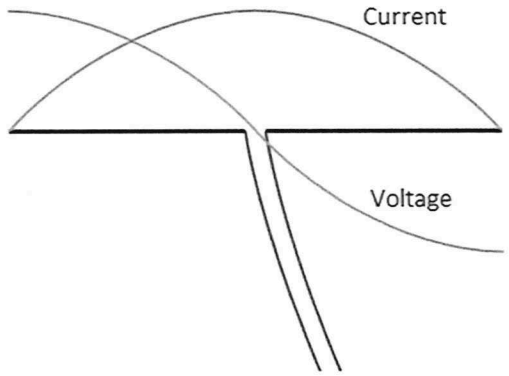


Figure 5 *The current and voltage waves through/across a half-wave dipole*

For reasons not explained here, the antenna is effectively a series resonant circuit. You may recall from your earlier AC studies that the current in a series resonant circuit is maximum at the resonant frequency. Given the current in a half-wave dipole is maximum in the middle, it must mean the middle is the dipole's resonant point.

You may also recall from your AC studies that, at resonance, the circuit's impedance is not only minimum (giving maximum current) but is also purely resistive.

So, the middle of the dipole is neither capacitive nor inductive. This matters because another thing you would have learnt from your AC studies is that capacitors and inductors don't dissipate power. Instead, they store it temporarily and return it to the source at another time (or swap the energy between each other) which is the opposite of what's needed for an antenna.

Why the half-wave dipole is an efficient design (continued)
Only resistance dissipates power and, as mentioned above, that's what the half-wave dipole looks like in the middle (and the resistance is 73Ω). So, by using a half-wave dipole **and** by connecting the signal to the middle, none of the energy sent to the antenna is returned to the source (ideally) and all of it is converted to an EM wave and propagated (ideally).

(In point of fact, as electrical signals on conductors travel at 95% of the speed of light, the ideal length of a half-wave dipole is 5% smaller than the calculations you performed earlier. However, this can be ignored for an introductory subject in communications.)

The quarter-wave monopole
The quarter-wave monopole (also known as the Marconi antenna) is one of the most widely used antenna designs for transmitting signals at frequencies in the MF band and the lower end of the HF band. Recall that the half-wave dipole is impractically long at, and below, these frequencies.

As its name implies, the Marconi antenna has only one conductor that is a quarter of a wavelength long. Importantly, it must be mounted vertically as the design uses the ground to implement the other quarter-wave conductor by producing an "image" of it and effectively forming a half-wave dipole. This is shown in Figure 6 below.

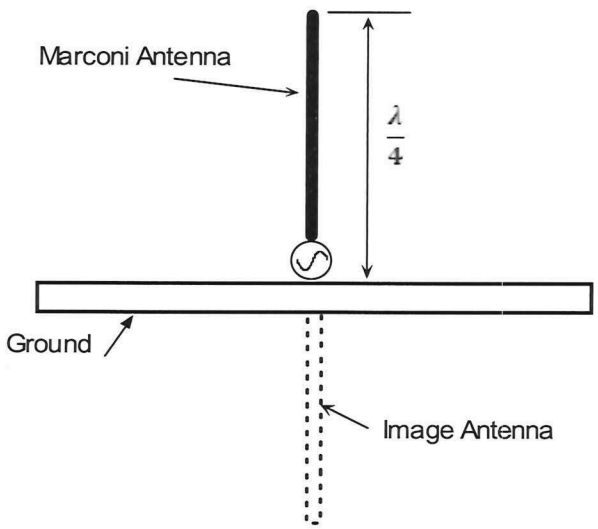


Figure 6 *The quarter-wave Marconi antenna*

An interesting fact about the Marconi antenna – Not tested
If the ground on which the Marconi antenna is to be located is not ideal, it can be improved by burying heaving copper conductors radiating away from the antenna's base (parallel to the surface) at 15° intervals.

Where the Marconi antenna is used for VHF and UHF applications, the ground plane can be replaced by a metallic conductor called a *counterpoise*. The roof of a vehicle can act as a counterpoise for mobile VHF and UHF communications.

Interestingly, the antenna in mobile phones is a type of monopole antenna.

Half-wave dipole antenna radiation pattern

The half-wave dipole antenna is said to be "directional" because it radiates energy in some directions but not all. We can consider its radiation pattern in one of two ways. The first is shown in Figure 7 below.

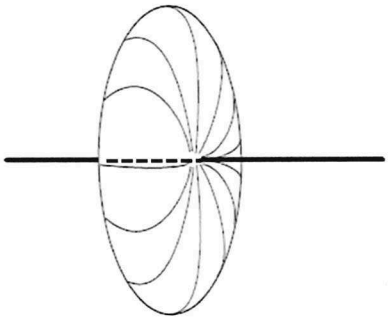


Figure 7 *The 3-dimensional radiation pattern of a half-wave dipole antenna*

There are a few points to make here about what Figure 7 shows us. First, the radiation pattern is donut-shaped but with a very small hole in the centre. Second, the antenna's radiation is maximum at 90° to its centre (in all directions) but zero along the axis of its conductors. Third, the actual emission of EM waves from the antenna is maximum at the antenna's centre where the current is maximum but decreases along its length. So, little if any energy is radiated from the antenna's ends in any direction. Fourth, close to the antenna, the wavefront of E and H fields is curved. However, at distance they straighten out.

If we consider a cross-section of the half-wave dipole's EM radiation in only two dimensions then we obtain the radiation patterns in Figures 8 and 9 (below). Figure 8 shows the cross-sectional radiation pattern looking at it from one of the dipole's ends (and hence the antenna is just a dot on the diagram). As you can see, energy is radiated in all directions in this plane.

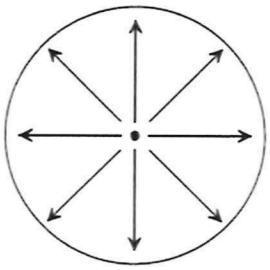


Figure 8 *The polar representation of a half-wave dipole antenna's radiation pattern from one of its ends*

Figure 9 shows the cross-sectional radiation pattern of the same antenna looking at it from any one of its sides.

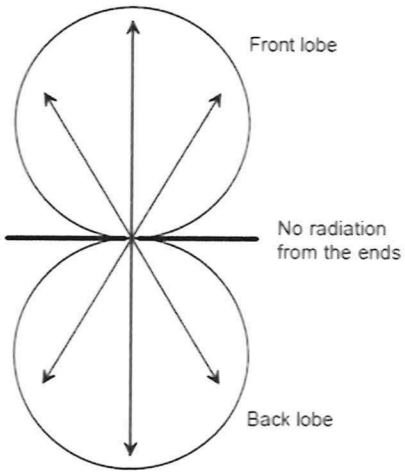


Figure 9 *The polar representation of a half-wave dipole antenna's radiation pattern looking at it from one of its sides*

The decision about which radiation pattern to consider depends on the antenna's polarisation. If the antenna is mounted vertically (and so the antenna is vertically polarised), the radiation pattern in Figure 8 is more relevant in terms of its "directivity" (more later in the chapter). If the antenna is mounted horizontally (and so it's horizontally polarised), the radiation pattern in Figure 9 is more relevant.

Figure 9 shows that, in this plane, the antenna's radiation pattern has two circular projections of radiated energy that are called *lobes*. These have been labelled the front and back lobe but this is arbitrary for Figure 9 because the lobes are identical.

Often, there are more than two lobes and just as often the lobes are not the same size or shape, in which case the lobe that radiates the most energy is considered to be the front lobe. An example is shown in Figure 10.

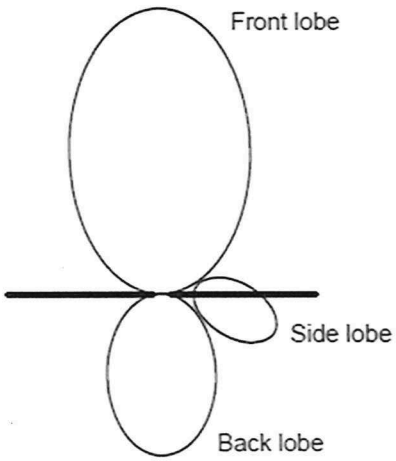


Figure 10 *An antenna's radiation pattern having three lobes of different size and shape*

The shape of the lobes depends on the proximity, shape and size of nearby objects including the ground. This fact is taken advantage of to produce specific radiation patterns for particular applications (more later).

The method of representing antenna radiation patterns in Figures 8 and 9 is widely used. It's called a polar diagram because it's drawn on polar graph paper so that useful quantitative information can be shown. An example is shown in Figure 11 below.

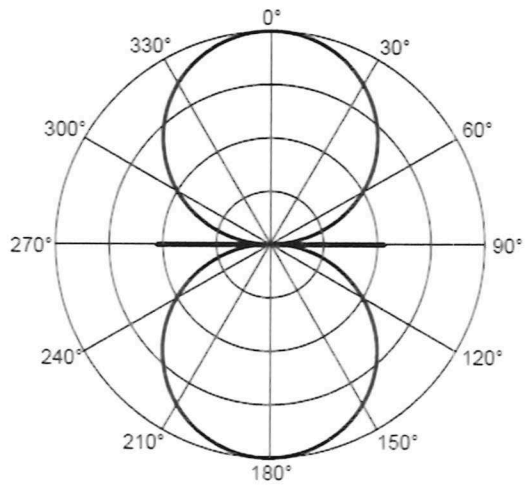


Figure 11 *A half-wave dipole antenna's radiation pattern shown on polar graph paper*

As you can see, the rotational axis is specified in degrees. The radial axis isn't quantified here because this can be done using one of several methods with a common one being power gain expressed in decibels relative to an isotropic source (and hence, dBi)

Antenna directivity

As its name implies, the directivity of an antenna is a term used to describe its ability to radiate more energy in one or more directions at the expense of other directions (or to receive more energy in one or more directions at the expense of other directions).

In fact, all antennas are directional to some extent because no antenna radiates equal amounts of energy in absolutely every direction. For some applications (like mobile phones) antenna directivity is undesirable and efforts are made to reduce it.

However, there are many situations in which there is a need radiate antenna energy in one or more directions in preference to others. For example, there's little point in sending broadcast radio and television signals out over the Tasman Sea. Apart from a few boaties, there are not many people off Sydney's coast to receive them. In these situations the antenna's directivity needs to be enhanced but in a particular way. There are several ways of achieving this (more later).

Antenna gain

One of the most important antenna parameters is gain. However, the term "gain" here is not being used in that classic sense where an output is compared to an input (eg Av = Vout/Vin). Instead, it's the increase in power radiated or received in a given direction (usually the front lobe) of a directional antenna compared to a reference. When the reference is an isotropic source (which is a theoretical antenna that radiates energy in all directions equally), the gain is expressed in dBi. An example of an antenna with a gain of 7dBi is shown in Figure 12 below.

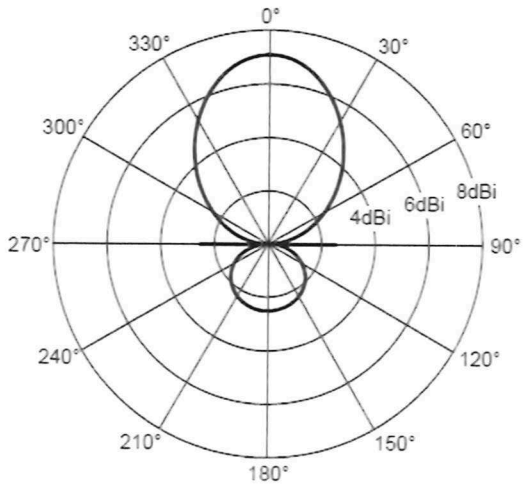


Figure 12 *The polar graph of an antenna with 7dBi gain*

Figure 12 can now be used to deepen your understanding of what polar graphs tell you. First, despite appearances, the graph is not showing the physical limits that the radiated EM energy reaches. This cannot be the case when you think about it. Once EM energy is radiated it travels indefinitely getting smaller as it goes until it strikes an object that absorbs it (or falls into a black hole but that's another story). So, as the gain at the top of the front lobe in the Figure 12 is 7dBi, the EM energy at every point along this line reaching back to the antenna and forward to infinity (and beyond!) is 7dB greater than the energy at the same points if the antenna was an isotropic source. This is highlighted in Figure 13 on the next page.

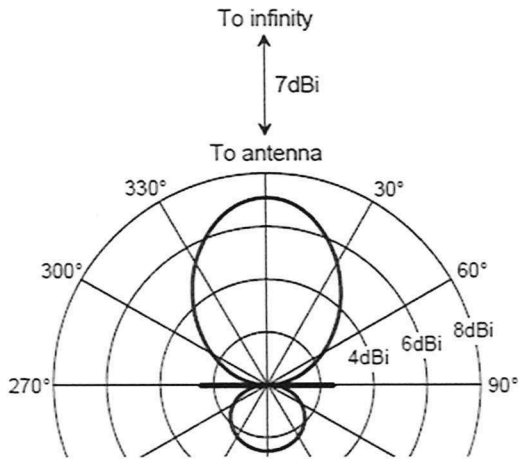


Figure 13

Second, although the antenna that Figure 12 represents has a front lobe gain of 7dBi, this is a maximum value and in one direction. As you move either side of this line (ie left or right) the gain goes down. This is highlighted for just one angle in Figure 14 below.

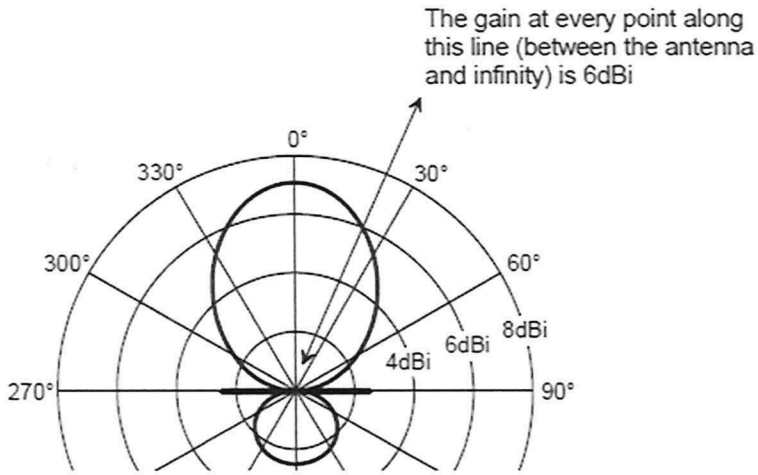


Figure 14

Importantly, this means even as a receiver moves around inside the antenna's lobe the EM energy it picks-up from the transmitter varies. [We don't tend to notice this though because receivers usually have automatic gain control (AGC) that compensates for changes in the field strength of signals (more later).]

The Marconi antenna radiation pattern

As the Marconi antenna is effectively a half-wave dipole (by using ground or counterpoise to create an image antenna) its radiation pattern is donut-shaped with a small centre hole like Figure 7. But to achieve the reflection in ground, the Marconi antenna must be mounted vertically as shown in Figure 15 below.

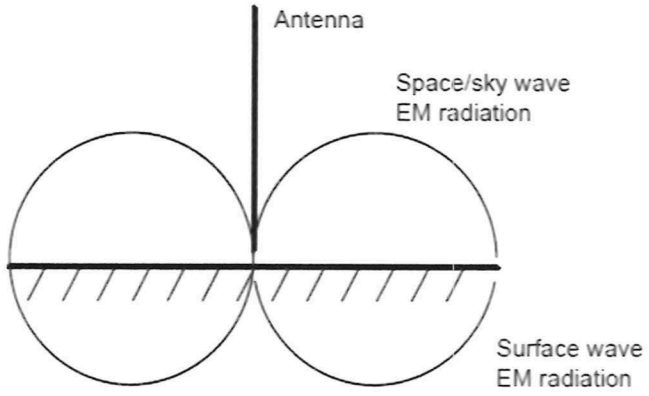


Figure 15 *The EM radiation of a Marconi antenna*

Observed from above, its geographic radiation pattern is shown in Figure 16 below with rotational axes removed for clarity.

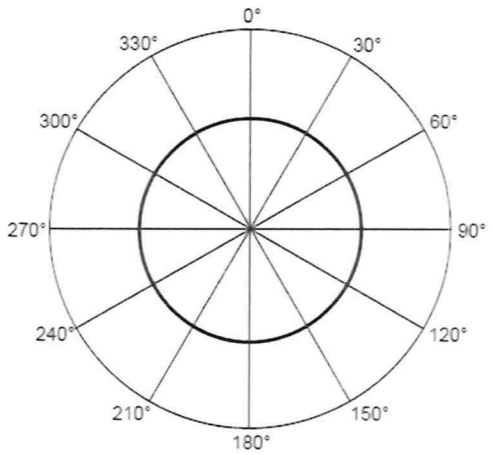


Figure 16 *The EM radiation pattern of a Marconi antenna*

The shape of the radiation pattern varies somewhat with the length of the antenna. If the antenna's length is increased, the donut shape flattens out increasing the energy projected at ground level at the expense of energy transmitted into the sky which is a good thing for AM broadcast radio. This is shown in Figure 17 below.

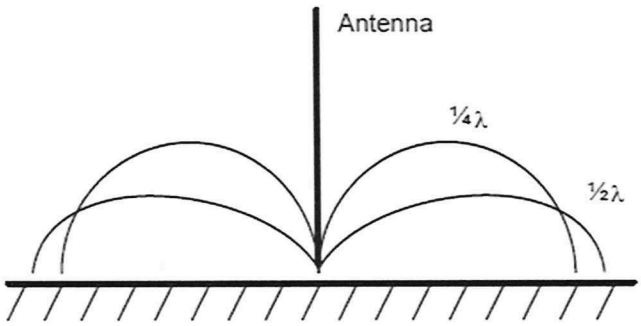


Figure 17 *The longer the Marconi antenna, the more its donut-shaped pattern flattens out*

The three-element Yagi antenna

A classic half-wave dipole antenna is the Yagi-Uda antenna (invented by two Japanese scientists in 1926). This antenna is used extensively for television reception and consists of three or more elements.

The main element is the half-wave dipole and is called the *driven* element (a reference to when the antenna is being used to transmit signals). The other elements are said to be "parasitic" because the absorb EM radiation and re-radiate it, creating an interference pattern in the EM waves that makes the dipole more directional (for reasons not explained here but you can explore this for yourself if you search online for EM wave interference patterns and directivity).

The element at the rear of the antenna is called the *reflector* and is slightly longer than the driven element at 0.55λ. The element at the front of the antenna is called the *director* and is slightly shorter than the driven element at 0.45λ. The elements are spaced 0.1λ apart. There is usually only one reflector but there can be more than one director.

The arrangement of the three-element Yagi together with key dimensions is shown in Figure 18 below. You're expected to be able to use this information in an assessment to design a three-element Yagi antenna optimised for the transmission/reception of a particular carrier frequency. However, you would be given the dimensions indicated below.

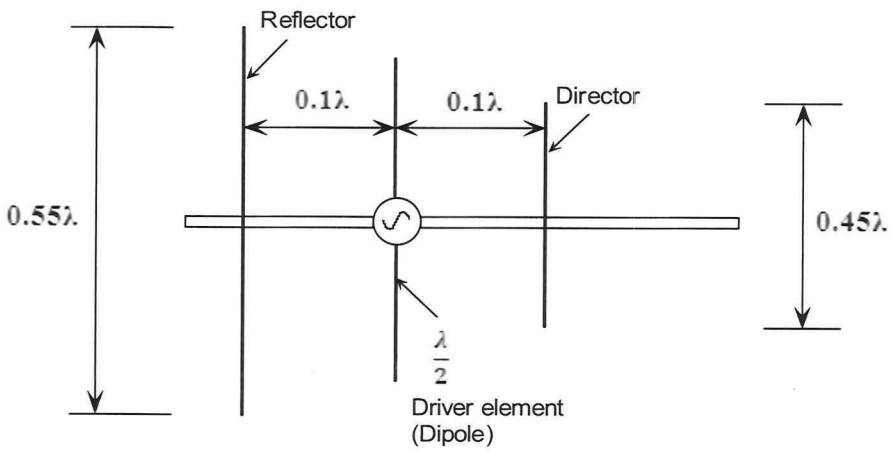


Figure 18 *The three-element Yagi antenna and approximate dimensions*

The radiation pattern for a three-element Yagi antenna is shown in Figure 19.

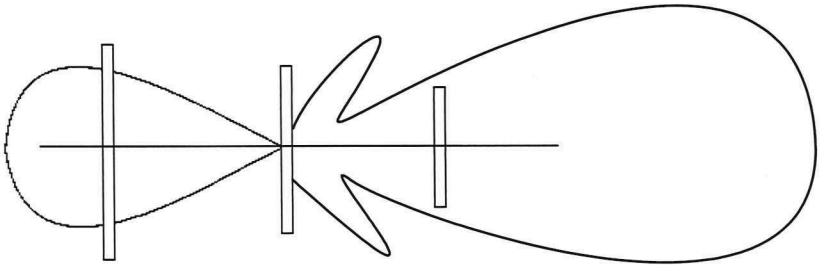


Figure 19 *The radiation pattern for a three-element Yagi antenna*

The three-element Yagi has a front lobe gain of about 7dBi. Apart from being better able to pick-up weaker signals, the directivity reduces the relative size of received reflected space waves at most angles which could cause reception problems.

Better directivity and hence higher gains (up to 16dBi) can be achieved with additional directors. Examples of antennas with more directors are shown in Figure 20 below.

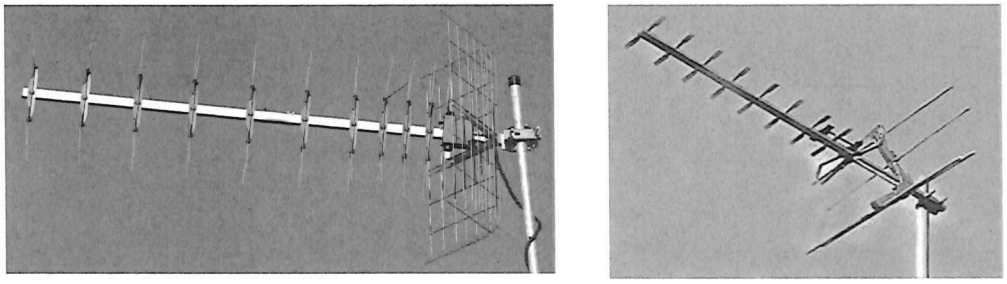


Figure 20 *Yagi antennas with multiple directors*

RG59 or RG6 coaxial cables are used for television receptions systems with RG6 being better for longer runs due its superior shielding. At the television, the cable is terminated with a PAL connector (also called a Belling-Lee connector) as shown in Figure 21 below. If RG6 cable is used, the cable is terminated with an F-connector and so an adapter can be used to convert the connection to PAL.

Figure 21 *A PAL connector is used to terminate coaxial cable for domestic tv antenna systems*

Other antenna designs – Not tested

There are a number of other antennas designs including:

- Folded dipole which has a wider bandwidth than the half-wave dipole
- Rectangular (or panel) antenna and used on mobile phone towers
- Parabolic dish which is highly directional and used for point-to-point microwave communication links and for satellite communications
- Helical antenna which is highly directional and used for the uplink in satellite communications as well as to produce radio signals with circular polarisation.

Safety precautions when working on antennas

There are always risks associated in working on any electronics equipment that is switched on. However, high power (kW) transmitters are more dangerous. These units often use vacuum tubes that require very high voltages (in the order of kV) which can be lethal! Therefore, always ensure that all high voltage supplies have been switched off before working on any circuits in the transmitter cabinet. Remember too that capacitors can store lethal charge long after the power has been switched off and so these should be appropriately discharged.
There are additional risks associated with the antenna(s) in communications systems. These include:

- Electrocution
- RF burns
- Working at heights

Electrocution

You're very unlikely to be electrocuted by a receiving antenna. The power of received signals is in the milli and micro watts and so these will not generate potential differences along the antenna sufficient to electrocute. That said, you could be unlucky enough to be holding the antenna while it is struck by lightning but that's another story.

Electrocution risks are mostly associated with transmission antennas and there are two possibilities both involving high transmission powers.

The first risk involves touching the antenna itself. If it's a Marconi antenna using a ground plane then touching the wrong point of the antenna while standing on ground can cause current to flow through your body to ground. Electrocution is also possible when touching a Marconi antenna and its counterpoise. And, electrocution is possible when touching two points of dipole or folded dipole antennas where the potential differences between those two points are sufficient high can also cause current to flow through your body.

The second electrocution risk relates to the electromagnetic radiation from high power transmission antennas. These may induce currents in nearby metal objects (either above or below ground) such as cables, plumbing, hand rails and ladders. If you come into contact with metal objects with induced currents inside you may receive an electric shock.

While electric shocks can be lethal in their own right, they can also cause other problems like involuntary spasms, falling into nearby machinery, and falling from your work area when working at heights (more on that later).

Take care to ensure that all high power transmission lines and antennas are unpowered before working on or near them. Always obey RF radiation warning signs and keep a safe distance from radiating antennas.

RF burns

There are two ways that burns can result from working near operating RF transmission equipment. The first is due to induced currents in your body and the second is caused by *dielectric heating*.

Your body is a conductor just like an antenna (though not as good, of course). Close contact with the sufficiently large electromagnetic radiation from RF transmission lines or antennas can result in burns due to the currents that the EM radiation can induce in your body.

Dielectric heating is a similar but different problem. Dielectric heating is the mechanism used to heat insulating objects (like food) in microwave ovens. When a human enters an RF field, their body becomes the dielectric in this process. If the power in the RF field exceeds 10mW per centimetre, the individual will have a noticeable rise in body temperature. The vital organs of the body and eyes are especially susceptible to dielectric heating.

RF burns are can be deep, penetrating, third-degree burns that can result in long-term injury and even death. Don't take RF burns lightly - Always seek treatment for any RF burn (or a shock for that matter) and report the incident to your supervisor so that appropriate action can be taken to correct the hazard.

For your safety, you must **not** stand directly in the path of RF radiating devices and you should avoid looking directly into devices radiating RF energy.

Before you work on an antenna, ensure that it is tagged-out properly at the switchboard and the switchboard locked to prevent it from being turned on. Always follow the correct safety procedures for working at heights.

Working at heights

The risk associated with working at heights is obvious – You can fall. Safework Australia has some useful information on managing these risks that can be found at:

https://www.safeworkaustralia.gov.au/heights

There are also industry standards governing the practice which can be found at:

https://www.waha.org.au/standards/

Importantly, you should really undertake a course on working at heights delivered by TAFE or one of the many private providers in the field. For an example, see:

https://www.pinnaclesafety.com.au/courses/height-safety/work-safely-at-heights-online-initial/nsw

Review questions

Answer these questions to check your understanding of what you have learnt from this chapter. Doing this will help you prepare for the knowledge assessment.

1. What is the basic construction of a dipole antenna?
2. Dipoles are efficient antennas but the Marconi antenna is a monopole antenna. Does this mean that it's much less efficient? Explain your answer.
3. What is the primary type of EM wave propagation used by: a) the half-wave dipole; and b) the Marconi antenna?
4. What does it mean when an antenna is said to be "horizontally polarised"?
5. Should a receiving antenna be in the same or opposite polarisation as the transmitting antenna? Explain your answer.
6. Calculate the length of a half-wave dipole antenna for 182.25MHz.
7. Sketch the radiation pattern for a horizontally polarised dipole antenna viewed from above.
8. Sketch the radiation pattern for a Marconi antenna.
9. Calculate the length of a quarter-wave monopole (ie Marconi) antenna at 1,269kHz.
10. Describe the purpose of the *director* and *reflector* in a three-terminal Yagi antenna.
11. Design a Yagi antenna for transmitting or receiving signals on a 198.5MHz carrier. You'll need to calculate the length of each of the elements and the spacing between them, and sketch the final design.
12. State the risks of working on or near antennas.
13. Briefly explain one of the two ways in which you can be electrocuted by touching an antenna.
14. Briefly explain how to minimise the risks when working on or near antennas.

Student notes

Section 4 Amplitude modulation and demodulation

This section aims to develop your understanding of the principles of amplitude modulation and transmission, how to perform AM modulation index and sideband power calculations, and your understanding of the principles of amplitude demodulation.

At the end of this section you should be able to:

a) Explain the general principle of amplitude modulation.

b) Sketch an AM signal in the time domain for a sinusoidal carrier that is amplitude modulated by a sinusoidal message and indicate on the drawing the signal's upper and lower envelopes.

c) Explain why AM is called is *double sideband full carrier* (DSBFC).

d) Sketch an AM signal in the frequency domain for a sinusoidal carrier that is amplitude modulated by a sinusoidal or complex message and indicate on the drawing the signal's upper and lower sidebands.

e) Calculate the frequency of all significant sinewaves in an AM signal for a sinusoidal carrier modulated by a sinusoidal message given all frequencies.

f) Calculate the lowest and highest frequencies in an AM signal given the carrier frequency and the highest frequency in a complex message.

g) Calculate the bandwidth of an AM signal given the frequency of a sinusoidal message signal or the highest frequency in a complex message signal.

h) Explain the connection between an AM system's allowable bandwidth and the fidelity of the recovered message.

i) Define modulation index as it applies to AM signals.

j) Draw examples of AM signals in the time domain for different levels of modulation, including overmodulation.

k) Calculate an AM signal's modulation index given either: a) the carrier and message voltages; or b) the AM signal's time domain minimum and maximum voltages.

l) Define overmodulation and state two problems that it can cause.

m) Calculate the total power in an AM signal and/or the power in an AM signal's sidebands given the carrier's unmodulated power and the signal's modulation index.

n) State the key difference between AM and ASK.

o) Give the full name of the AM modulation scheme represented by the initials SSBSC.

p) Briefly describe (including with the use of frequency domain representations) the difference between DSBFC signals and SSBSC signals.

q) State two advantages of SSBSC over DSBFC.

r) Sketch the circuit diagram of an envelope detector and briefly describe its operation.

s) Explain why the envelope detector cannot be used to recover the message from SSBSC signals.

t) Explain why DSBFC AM communications is affected by noise superimposed on the AM signal.

u) Practise using an oscilloscope to observe the envelope of an AM signal (modulated by a sinusoidal message) and determine its modulation index by measurement.

v) Practise using an oscilloscope to observe the envelope of an AM signal (modulated by a sinusoidal message) when it is over-modulated.

w) Practise using an oscilloscope to observe the envelope of an AM signal modulated by a complex message signal such as speech.

Introduction

You have already learnt that the relatively low frequency of message signals makes them unsuitable for direct transmission. For this reason, the message signal (whether it's speech, music, video or data) is normally "encoded" onto a higher frequency carrier (usually sinusoidal) in a process called "modulation".

Two other terms used to describe modulation include the message being "impressed" or "superimposed" onto the carrier.

Useful info – Not tested
All three words (encode, impressed and superimposed) are metaphors but they're all helpful in understanding modulation. It would be worth your while looking-up the definitions of these three words to help deepen your understanding of modulation.)

The two main modulation methods used in broadcast radio are amplitude modulation (AM) and frequency modulation (FM). AM and FM are used in other communications applications too, as are phase modulation (PM), quadrature amplitude modulation (QAM), and at least a dozen others. As this is an introductory unit on communications fundamentals, the treatment of modulation is limited here to just AM and FM. There is an elective unit on communications that you can take if you would like to know more about some of the other modulation schemes.

For more useful info – Not tested
The Wikipedia pages on AM and FM (and the other modulation schemes) are quite good for:

- Animations demonstrating what the modulated signals look like
- The history of their invention and early use
- Modern applications
- A different perspective on the basic theory
- More advanced theory if you're interested to know more

You should definitely visit these pages when you have the time.

Amplitude modulation

With amplitude modulation (AM), the amplitude of the carrier signal is varied by the message signal's amplitude. To be more specific, the carrier's instantaneous amplitude is directly proportional to the message signal's instantaneous amplitude.

Importantly, this is not achieved by simply combining the message signal and carrier signal using a linear device such as a resistor. This would simply add the two signals together and so no encoding takes place.

To explain, suppose the message is a 1kHz sinewave and the carrier is a 20kHz* sinewave. Simply adding the two signals would give a waveform with an interesting shape. However, spectrally, it would still just consist of the two original sinewaves – The message will not have been encoded. Moreover, with the message signal still present in its baseband form, all the problems of transmitting message signals discussed previously are still applicable.

> *This is too low a carrier frequency for most communications systems but it makes it easy to interpret the resultant in the Excel spreadsheet that your author uses in class.

Instead, AM is produced by "mixing" the message and the carrier using a non-liner device such as a diode. Figure 1 below shows the resulting AM signal for a given message signal (*fm*) and a given carrier signal (*fc*).

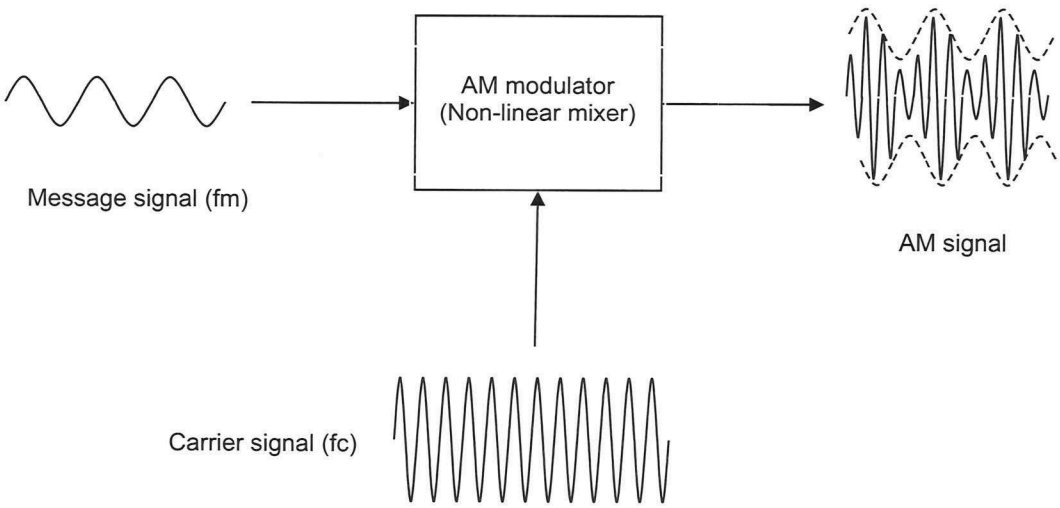


Figure 1 *The message and carrier signals combine (are multiplied) in the modulator/mixer to produce the AM signal*

To give you a clearer idea of the effect of the message on the on the carrier, Figure 2 below shows the message and carrier time-coincident with each other. The message starts at 0Vpp and so the carrier's peak-to-peak voltage is unchanged. However, when the message signal goes positive, the carrier voltage becomes proportionally bigger (relative to the unmodulated carrier's peak-to-peak voltage). And when the message signal goes negative, the carrier voltage becomes proportionally smaller (relative to the unmodulated carrier's peak-to-peak voltage).

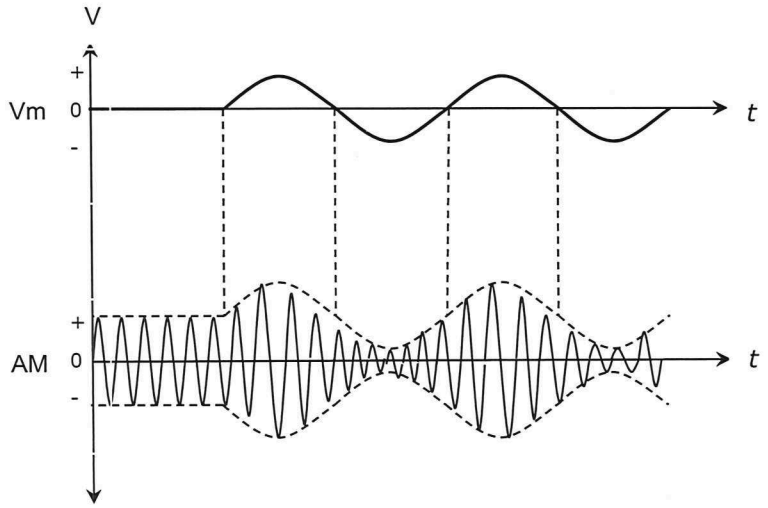


Figure 2 *The message and carrier signals time-coincident with each other*

Using dashed lines to join all the positive peaks to each other and all the negative peaks to each other identifies the AM signal's "envelope".

If you look closely, you'll notice that the upper envelope (the line joining all the positive peaks) is the same shape and frequency as the message. And, the lower envelope (the line joining all of the negative peaks) is the same shape and frequency as the message but inverted. These are the distinctive features that identify a modulated signal as AM. (That is, other modulation schemes using the same two signals produce a waveshape that is completely different.)

It can be proven mathematically (not shown here) that when a sinusoidal message signal (f_m) and sinusoidal carrier signal (f_c) are mixed using a non-linear device, a DC voltage is produced together with a collection of sinewaves at the following frequencies:

- the message's frequency (f_m)
- the carrier's frequency (f_c)
- the sum of the two frequencies ($f_c + f_m$)
- the difference of the two frequencies ($f_c - f_m$)
- whole number multiples of the message frequency ($2f_m$, $3f_m$ and so on to ∞)
- whole number multiples of the carrier frequency ($2f_c$, $3f_c$ and so on to ∞)

Some of these spectral components are shown in the frequency domain graph in Figure 3.

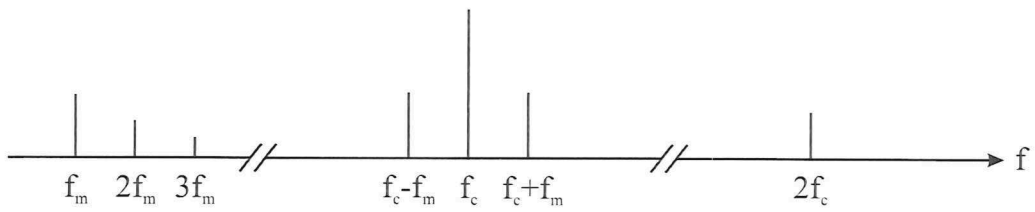


Figure 3 *Spectral composition of non-linear mixing of two sinewaves*

Most of these frequency components in the AM signal's spectrum are not significant in terms of their size and have only been mentioned here for completeness. The small number of these components that are significant but are not needed are removed by filtering prior to transmission.

When "mixing" is adding or multiplying– Not tested
"Mixing" is a term that is used by both audio and RF technicians/engineers. When audio people talk about mixing, they're referring to a process which produces a resultant that is the simple mathematical sum of the signals' voltages and no new sinewaves are produced. But when RF people talk about mixing, they're referring to a process which produces a resultant that is the mathematical multiplication of the signals and new sinewaves are produced per the example above.

Spectral composition of AM and the sideband frequencies

As the previous discussion explains, the non-linear combination of a sinusoidal message and sinusoidal carrier produces a lot of sinewaves. However, only the sinewaves shown in Figure 4 below are necessary for an AM signal to be faithfully demodulated at the receiver.

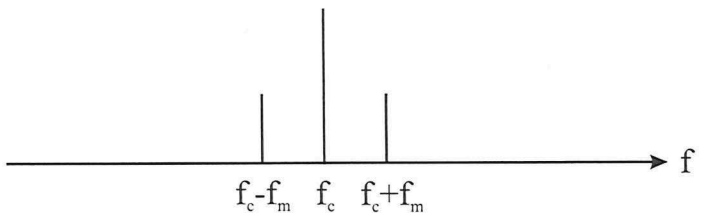


Figure 4 *AM components shown in the frequency domain*

In other words, the spectral composition of an AM signal for a sinusoidal message (f_m) and sinusoidal carrier (f_c) is:

i) a sinewave at the carrier's frequency (f_c)

ii) a sinewave that is the difference of the two frequencies ($f_c - f_m$) and is called the lower sideband (f_{LSB})

iii) a sinewave that is the sum of the two frequencies ($f_c + f_m$) and is called the upper sideband (f_{USB})

Doing an example with numbers, suppose a 1MHz sinusoidal carrier is amplitude modulated by a 5kHz sinusoidal message signal. The resulting AM signal will consist of three sinewaves. These are:

i) $f_C = 1\text{MHz}$

ii) $f_{LSB} = f_c - f_m = 1\text{MHz} - 5\text{kHz} = 0.995\text{MHz}$ (or 995kHz)

iii) $f_{USB} = f_c + f_m = 1\text{MHz} + 5\text{kHz} = 1.005\text{MHz}$ (1,005kHz).

It's for this reason that this type of AM signal is called "double sideband full carrier" which is often abbreviated to DSBFC. DSBFC AM is used for AM broadcast radio. Other forms of AM exist and one of these are described later.

The sidebands are crucial because they both individually contain the original information in the message. That being the case, you may understand now why the term "encoding" is a good one to use when talking about modulation. The original information in the message (the 5kHz sinewave) has been encoded into two other sinewaves, one at 995kHz and one at 1,005kHz.

Importantly, none of the information in the message is present in the carrier component and so, theoretically, it's not needed. However, the carrier is essential to the operation of the simple circuit that is widely used to demodulate AM (more later).

Amplitude modulation using pure multiplication – Not tested
If you were to multiply the message (*fm*) and the carrier (*fc*) using a mathematically pure multiplication process (eg by using an analog multiplier circuit instead of a diode circuit), the resultant consists of only the sinewaves required for the AM signal (ie fc - fm, fc and fc + fm).

Discrete component analog multiplier circuits are challenging to implement so mixing, together with filtering, has historically been preferred for generating AM signals because it was much simpler. This method is also better suited to the circuits needed to amplify AM signals sufficiently for broadcast radio.

Integrated circuit multipliers are available nowadays (like the AD633) and these can be used for generating AM signals for low power transmission applications.

AM signal bandwidth

The total transmission bandwidth of an AM signal is the difference between the upper and lower sideband frequencies. Using the figures in the example of an AM signal above, we can see that the bandwidth is 1,005kHz – 995kHz which is 10kHz.

Notice that this is twice the message frequency and this gives us the general equation for the BW of AM signals from sinusoidal messages:

$$\boxed{BW = 2 \times f_m}$$

The spectral composition of AM for complex message signals

So far, for the purposes of making AM easy to explain and understand, the message signal has been a single sinewave. This is seldom the case and is definitely not the case for AM broadcast radio where the message signals are speech and music.

Speech, music, video and data are very complex signals comprising a range (or band) of sinewaves at many frequencies. For example, the fundamental frequencies produced by human speech sounds range from hundreds of Hertz to thousands of Hertz (the figures varies depending on how high or low the person talks). Obviously, our speech sounds will include harmonics at higher frequencies too, but not much beyond 8kHz for men and 17kHz for women.

Just to put a convenient number on it, let's say that a message signal includes significant sinewaves at frequencies from 200Hz to 5kHz. On a frequency domain graph, the message could be represented as shown in Figure 5 below.

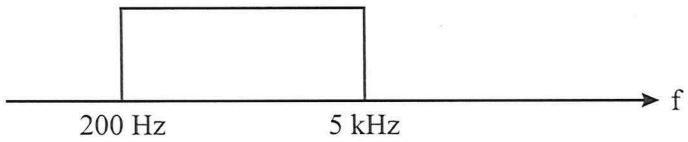


Figure 5 *A complex message including many sinewaves at frequencies from 200Hz to 5kHz*

Importantly, if a complex waveform such as Figure 4 is used to amplitude modulate a sinusoidal carrier, a pair of sinewaves is generated in the AM signal for every sinewave in the message, one in the AM signal's lower sideband and one in its upper sideband. Their frequencies are calculated the same way as on the previous page.

This means that, in the frequency domain, the AM signal of a 1MHz carrier modulated by the message in Figure 5 looks Figure 6 below.

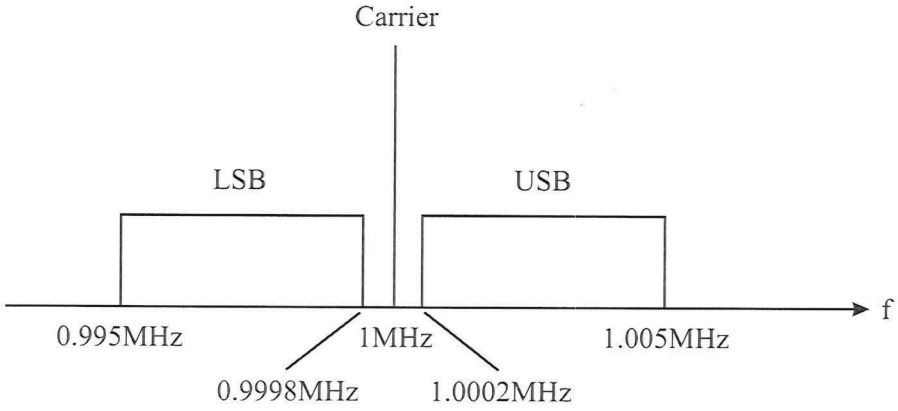


Figure 6 *AM signal on 1MHz with a complex message consisting of sinewaves at frequencies from 200Hz to 5kHz*

The bandwidth of this AM signal is found by subtracting the lowest frequency in the AM signal from the highest frequency. So, 1.005MHz – 0.995MHz = 10kHz.

You'll notice that this is the same for the bandwidth of a 1MHz carrier modulated by a single 5kHz sinewave a couple of pages ago. So, complex messages don't change the bandwidth equation. However, the equation needs a little refining to indicate that it is calculated using the highest frequency in the message. That is:

$$\boxed{BW = 2 \times f_{m(max)}}$$

Communication system bandwidth and fidelity

The telephone system was originally designed to transmit the speech frequencies from 300Hz to 3.4kHz. Even though this doesn't permit the transmission of all frequencies in speech sounds, you'll know from experience that it's adequate for intelligible conversation and the person's voice will sound very much like it does in face-to-face conversation. In other words, the reproduced speech out of your phone has a reasonable degree of "fidelity" or faithfulness to the original.

But, you'll also know that if you're talking on the phone and you're put on hold and made to listen to music, the music will not sound nearly as good as when you're listening to it on the radio or using an MP3 player. That's because musical sounds include many more fundamental frequencies and harmonics at higher frequencies than speech typically does and a significant number of them are at frequencies outside the bandwidth of the telephone system's design.

Listening to music on AM radio will sound better than listening to music over the phone because AM broadcast radio has be designed to permit the transmission of a greater frequency range of complex message frequencies.

Listening to music on FM radio will sound better than listening to music both on AM radio and over the phone because it has been designed to permit the transmission of an even greater frequency range of complex message frequencies.

Regardless of whether the message is speech, music, video or data, the greater the frequency range of complex message frequencies that are permitted to be transmitted, the better the system's fidelity.

Let's consider this issue with AM broadcast radio in Australia. The bandwidth allocation for each radio station is 17kHz. If we use the bandwidth equation to work backwards, we find that the highest frequency in the message that's permitted to be transmitted is 8.5kHz (ie 17kHz ÷ 2). That's why listening to music on AM radio sounds better than listening to music over the phone.

We can't do the calculations for FM radio just now because FM hasn't been explained. But it's a fact that FM broadcast radio in Australia permits the transmission of message frequencies up to 15kHz. That's why listening to music on FM radio sounds much better than listening to music both on AM radio and over the phone.

Could AM radio have been designed to permit the transmission of message frequencies up to 15kHz? Absolutely! But, the bandwidth for each station or channel would be 30kHz and so that would permit just over half as many radio stations in a given band than a system with a bandwidth allocation of 17kHz.

Does that matter when we have FM radio as well as AM radio? Probably not any more but for a good 50 years or more, AM radio was the only broadcast radio system available and so economical use of that band was essential. The bandwidth allocation was a balance between the number of permissible stations and the fidelity of the received signal from each station.

For communication systems where the expectation is that the message is only going to be speech, the bandwidth for each channel is made much more narrow than for broadcast radio so that more channels (phone calls) can operate on a band. We see this in telephony where the bandwidth is only 3.1kHz (ie 3.4kHz – 300Hz).

Modulation Index

The term "modulation index" (sometimes called percentage modulation) is a measure of how deeply the carrier is modulated by the message signal.

To explain, the drawing of an AM signal earlier in this chapter shows how it would look for a certain peak-to-peak message voltage (relative to a certain unmodulated carrier's peak-to-peak voltage). If the size of message changed relative to the carrier, the dimensions of the AM signal's envelope would also change. Figure 7 below shows two examples of AM signals with different depths of modulation. They indicate that, for the same sized carrier, the AM signal on the right has a bigger message.

Figure 7 *AM signals with different depths of modulation*

As the AM signal's depth of modulation is a function of the message's amplitude relative to the carrier's amplitude, its modulation index (m) is simply the ratio of the two. So, if the amplitudes of the message and carrier are known, the modulation index can be calculated using:

$$m = \frac{V_{message}}{V_{carrier}}$$

As the amplitude of the message signal increases relative to the carrier's amplitude, so does the modulation index.

The modulation index is oftentimes expressed as a percentage and the equation for this is:

$$m = \frac{V_{message}}{V_{carrier}} \times 100$$

You're welcome to express modulation index as either a ratio or a percentage. However, the ratio figure must be used when the modulation index is needed for power calculations (more later).

When you don't know the size of the message and carrier signals, and you can't measure them, there's an alternative method of determining the modulation index that relies on the AM signal itself.

The key dimensions of the AM signal needed for this calculation are its minimum peak-to-peak voltage and its maximum peak-to-peak voltage. These are shown as dimensions A and B respectively in Figure 8 below. (Note: The letters "P" and "Q" are often used in place of "B" and "A" respectively.)

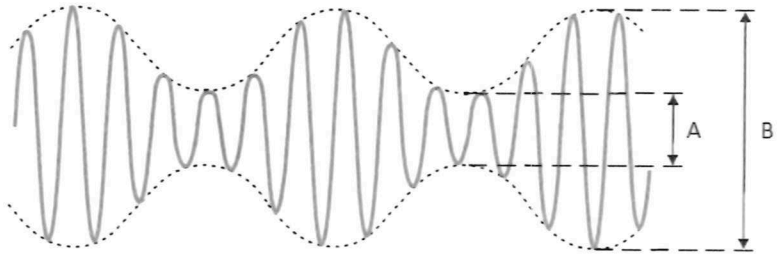


Figure 8 *The AM envelope and modulation index*

Importantly, the more deeply a carrier is modulated, the smaller dimension A becomes and the bigger dimension B becomes. Moreover, the difference between these dimensions has a proportional relationship to the amplitudes of the message and carrier such that it allows us to calculate the modulation index using the equation:

$$m = \frac{B - A}{B + A} \times 100$$

Let's do an example. Calculate the modulation index of an AM signal with a maximum voltage of 10Vpp and a minimum voltage of 2Vpp.

$$m = \frac{B - A}{B + A} \times 100$$

$$m = \frac{10V - 2V}{10V + 2V} \times 100$$

$$m = 66.7\%$$

The maximum percentage modulation is 100%. At 100% modulation, the AM waveform has a minimum peak-to-peak value for the A dimension and equal to zero. This is shown in Figure 9 below.

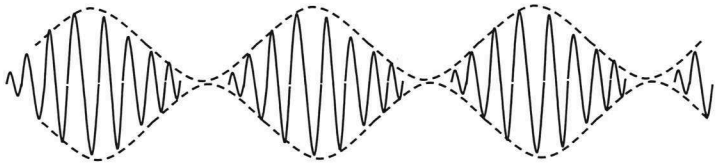


Figure 9 *An AM envelope with a modulation index of 100%*

Importantly, using the equations on page 75, it's notionally possible to obtain a modulation index that is greater than 1 (or 100%) but this becomes meaningless in terms of what the modulation index figure is used for (ie power calculations).

Overmodulation

Overmodulation occurs when the modulation index exceeds 100%. An over-modulated AM waveform is shown in Figure 10. Notice that the bottom peaks in upper envelope and the top peaks in the lower envelope overlap.

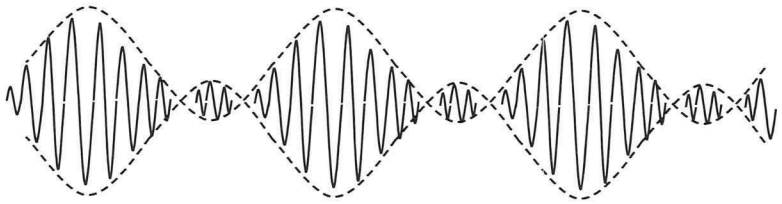


Figure 10 *An over modulated AM waveform*

This condition is undesirable for two reasons:

1) Where Class C amplifiers are used at the output stage of the transmitter (ie where "high level modulation" is used), overmodulation causes additional spectral components to be generated at frequencies just below the lower sideband and just above the upper sideband. That can cause the channel's allocated bandwidth to be exceeded and interfere with the signal of an adjacent channel, an effect called "sideband splatter".

2) Where the envelope detector circuit is used to demodulate the AM signal (more later) the fact that the envelopes are no-longer copies of the message renders the envelope detector incapable of recovering the message without significant distortion.

Amplitude shift keying

As already mentioned, the message signal in communications systems can be digital data. When data is used as the message in amplitude modulation, the system becomes known as *amplitude shift keying* (ASK).

Figure 11 below shows a simple squarewave being used to represent a message that is data. (This is done for visual simplicity as a true data message would produce a more visually complicated drawing that, if used, might detract from understanding.) Time-coincident with the message is the modulated carrier (the ASK signal).

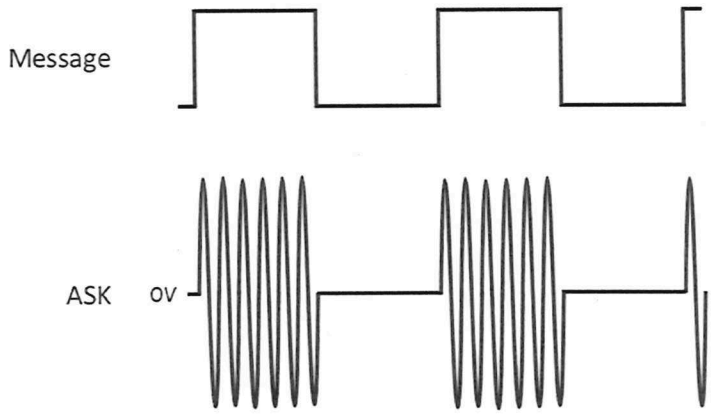


Figure 11 *A squarewave message time-coincident with the carrier that it is being used to amplitude modulate*

Line codes – Not tested

It's seldom the case, if ever, that the raw digital data message signal is used to directly modulate the carrier for any data in any modulation scheme. The data message signal is invariably converted to a "line code" first and then used to modulate the carrier. This is done for several reasons that you can explore for yourself online. Common line codes used include (but are not limited to); NRZ-L, Bi-Phase, RZ-AMI, and NRZ-M. These line codes are also known by other names.

Importantly, the carrier is typically 100% modulated and this is done to spread the AM signal's peaks to reduce the risk of noise corrupting the signal (more on noise and AM signals later). So, this means that the carrier voltage is reduced to 0Vpp when the data is logic-0.

Although ASK can be generated using non-linear mixing like AM generated using an analog message, Figure 10 suggests another method of modulation can be used to generate ASK. Simply using the data to switch on and off a transmitted would result in a waveform with an identical envelope. That is, the carrier is switched on when the message is a logic-1 and switched off when the message is logic-0.

A little history to explain the name – Not tested

"Keying" is a term that dates back to the earliest form of electrical communications; the telegraph. It's a reference to the spring-loaded normally-open switch (called a "key") that was used to tap out telegraph messages in Morse code.

As you probably know, Morse code makes use of two elements; dots and dashes. Pressing the Morse key for a short time transmitted a dot and pressing the key for a little longer transmitted a dash.

But actually, when you think about it, Morse code involves three elements because when a dot or a dash isn't being transmitted, no signal is being transmitted. This is called a "null" and it's as important an element in Morse code as the dots and dashes - Without the null, the dots and dashes can't be separated.

The term "keying" is borrowed and used as a name for many digital modulation techniques like ASK, FSK, BPSK, QPSK, DPSK, and so on.

This makes perfect sense when you look at the ASK signal Figure 10. When the message is logic-0, the AM signal is literally null. That said, as you will see with FSK, a signal is transmitted when the message is a logic-1 and a logic-0 and so "keying" essentially becomes metaphoric.

Power in an AM signal

The purpose of calculating the modulation index of an AM signal is to permit calculations of the power in the AM signal's sidebands. This is important because all of the information in the original message is contained in the AM signal's sidebands; no information is carried on its carrier. So, the more power in the sidebands, the better in terms of transmission distances.

Interesting fact – Not tested
As a rough guide, AM broadcast radio transmitters transmit powers from about 250W to 50kW or more. CB radio is generally limited to around 4W.

The total power in an AM (DSBFC) signal can be found using the equation:

$$P_t = P_c\left(1+\frac{m^2}{2}\right)$$

Where:
P_t = Total transmitted power (carrier and sidebands)
P_c = Carrier power
m = Modulation index (as a ratio)

Let's do an example. Calculate the total power transmitted when a transmitter's 2000W unmodulated carrier is modulated to a depth of 45%.

$$P_t = P_c\left(1+\frac{m^2}{2}\right)$$

$$P_t = 2000\left(1+\frac{0.45^2}{2}\right)$$

$$P_t = 2202.5W$$

As the carrier's power is constant, this means that the additional power in the AM signal must be the power in the sidebands (P_i) which can be calculated using: $P_i = P_t - P_c$.

So, for the example above, the combined power in the sidebands is 202.5W (proof: P_i = 2202.5W – 2000W = 202.5W).

(That power is split equally between the two sidebands but that's academic because AM demodulation "harvests" the energy in both sidebands to recover the message.)

Importantly, calculating the power in the sidebands is not just some academic exercise in number-crunching. To explain, let's do another example. This time, suppose a second transmitter's 500W unmodulated carrier is being modulated to a depth of 90%.

$$P_t = P_c\left(1 + \frac{m^2}{2}\right)$$

$$P_t = 500\left(1 + \frac{0.9^2}{2}\right)$$

$$P_t = 702.5W$$

And a quick calculation shows that the sideband power is 202.5W, the same as it was for the 2000W unmodulated carrier modulated to a depth of 45%.

Do you see the significance?

The two AM signals both have 202.5W in their sidebands and this is the energy that is harvested by the demodulator at the receiver. So both signals would sound equally loud (all other things being equal) but the second transmitter achieved that by transmitting a little under a third of the total power.

To think about this another way... Suppose our second transmitter used a 2000W unmodulated carrier like the first but with a modulation index of 90%. The same calculations would show that the sideband band power is 810W or four times the sideband power of the first example. (That said, the distance that the transmitted signal can theoretically travel is only double because of the inverse square law relationship of distance and signal strength.)

From these calculations, we can understand why AM transmission systems try to obtain a modulation index of 90% - 95% to maximise the power in the sidebands. That said, this must be achieved while preventing overmodulation which has the problems described earlier in the chapter.

Other types of AM

Single sideband suppressed carrier

If power in the carrier of a DSBFC signal is wasted, why transmit it at all? Moreover, if each of the sidebands contains all of the original information in the message, why send them both? These good questions lead to the development of another type of AM called "single sideband suppressed carrier" (SSBSC, or often just SSB).

As its name implies, SSBSC involves transmitting just one of the sidebands (either the upper or lower sideband, it doesn't matter which). Figure 12 illustrates upper sideband SSB transmission in the frequency domain for a complex message such as speech.

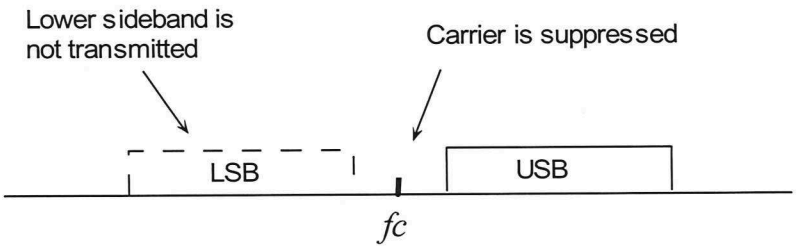


Figure 12 *Upper sideband SSBSC*

The main advantages of SSB are two-fold. First, as the carrier isn't transmitted, there is no wasted power and so the efficiency of SSB transmissions is 100%. Second, only half the channel bandwidth is needed for complex message signals compared to DSBFC AM. This permits twice as many channels (ie users) as AM for the same band of frequencies.

Broadcast AM radio uses DSBFC not SSBSC – Not tested

Why isn't SSBSC used for broadcast radio when it has two very significant advantages over DSBFC? The reason is historical. Demodulating SSB signals is a very technically challenging proposition compared to demodulating DSBFC and so more complex circuitry is needed. When broadcast radio was first devised, using SSBSC would have made domestic AM broadcast radio receivers much more expensive, undermining the economic viability of broadcast radio in the first instance - Fewer people could have afforded a receiver to be listening to the ads.

These days, the circuit that's used to demodulate SSB (called the *product detector*) can be bought fully implemented inside an IC for under $2! Interestingly, the product detector can also demodulate DSBFC AM.

Double sideband suppressed carrier – Not tested

Double sideband suppressed carrier (DSBSC) is another type of AM that doesn't transmit a carrier. That said, as its name implies, it does transmit both sidebands.

The benefits, or not, of DSBSC depend on what other version of AM that you're comparing it to. Compared to DSBFC AM, DSBSC is more efficient but requires the same bandwidth. Compared to SSBSC, DSBSC has the same efficiency but requires double the bandwidth.

[Wikipedia and another website suggest that DSBSC has an efficiency of 50% compared to SSBSC. Your author doesn't believe this to be correct because the necessary demodulation circuit for DSBSC (the product detector) can be shown mathematically to harvest the energy from both sidebands to recover the message. So, there's no wasted energy just like with SSB.]

While DSBSC is seldom used as a modulation scheme for analog communications systems in its own right, it is used to implement other features within in analog communications systems (such as stereo encoding in broadcast FM radio).

"Binary Phase Shift Keying" (BPSK) is a modulation scheme used for communications systems involving data for the message. Interestingly, the spectral composition of BPSK signals demonstrates that it's actually a form of DSBSC.

AM demodulation

The term "demodulation" (often referred to as "detection" for AM receivers) is the name for the process of recovering the information in the original message used to modulate an AM signal.

There are two types of AM demodulator: 1) The product detector (but they're not discussed in this Unit); and 2) The envelope detector (also known as the diode detector). Historically, the envelope detector has been the most widely used of the two circuits for low cost receivers due to its simplicity – It consists of just three components.

The circuit of an envelope detector is shown in Figure 13 below.

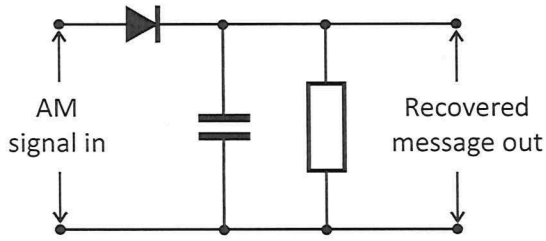


Figure 13 *The circuit of an envelope detector*

In addition to the circuit being simple, the explanation of how it works is simple too.

Suppose the circuit consisted of just a diode. And suppose we connected an AM signal to its input that is a sinusoidal carrier modulated by a sinusoidal message. The circuit's output would look like that in Figure 14 below. As you can see, the diode half-wave rectifies the AM signal and so one of the envelopes has been removed.

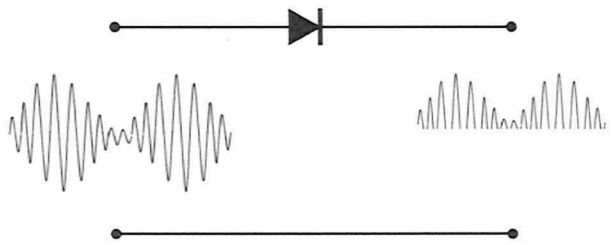


Figure 14

Now suppose the capacitor is added, it would charge up to the positive peaks on the rectified AM signal's way up but, because capacitors store charge, its potential difference would remain at a value near the highest peak as it slowly discharges. This is shown in Figure 15 below.

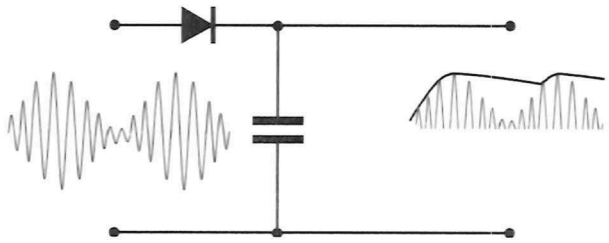


Figure 15

Finally, by including the resistor (with just the right value), the capacitor is allowed to discharge between the peaks so that its potential difference tracks the peaks reasonably closely. This is shown in Figure 16 below and, as you can see, the message has been recovered. This works because the AM signal's envelopes are the same shape as the original message used to modulate it.

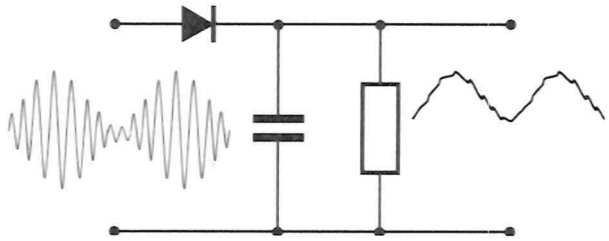


Figure 16 *Envelope detector with AM signal in and copy of message out*

There are some important and/or useful points that need to be made about this circuit.

1) You're not expected to be able to determine appropriate values for the capacitor and resistor for any assessment for this Unit.

2) As the drawing shows, the recovered message consists of a small ripple voltage at the carrier frequency. This is called "carrier feed-through". Even if this ripple is able to pass through the audio circuits (they may not have sufficient bandwidth because of the carrier's frequency), it'll not be heard by the listener because its frequency is inaudible to human hearing.

3) If the diode is reversed, the circuit detects the lower envelope instead of the upper envelope and so the recovered message would be inverted. Interestingly, this signal doesn't need to be inverted again by the audio circuits to correct that - Your hearing can't tell the difference between sounds and inverted versions of the same sounds.

4) The non-linear performance of the diode means that the recovered message is slightly distorted (with one website claiming that it introduces a THD of 5 – 10%). Although this is satisfactory for low cost radio systems, it's less than ideal for high-end hi-fi systems having much lower THD levels overall.

5) Germanium diodes (and more recently, Schottky diodes) are used for the envelope detector because of their lower forward voltage drop (about 0.2V to 0.3V) which permits demodulation of smaller signals. Think about it, if a regular diode is used, the AM signal needs to be a fair bit larger than 1.2Vpp for the envelope detector to work.

6) You can think of the operation of this circuit slightly differently. If we think of the half-wave rectified signal on the diode's output as a complex waveform (which it is), a spectral analysis of it would show that the sinewave with the lowest frequency is exactly the same frequency as the original message. In which case, we can think of the capacitor and resistor as providing a low-pass filter to filter out the high frequency carrier (and a bunch of other sinewaves too), leaving only the copy of the message to pass to the audio stages.

7) The circuit's output includes a DC voltage that is proportional to the size of the signal on its input. This DC voltage is often used in AM receivers to help implement a feature called "automatic gain control" (AGC) which is discussed later (in the treatment of superheterodyne receivers).

8) The envelope detector cannot recover single sideband (SSB) and double-sideband (DSB) signals without producing severe distortion. That's because the envelope of these two modulation schemes is not a copy of the original message.

AM and its susceptibility to noise introduced in the medium

For wireless communications like broadcast radio, probably the single biggest problem with DSBSFC AM is its susceptibility to noise that is introduced in the medium.

To explain, the medium is full of electromagnetic (EM) energy. Plenty of this is human made from the plethora of wireless communications taking place over the "air waves" (I prefer the term "free space"). And much of it is also caused by the power distribution system (emitted by the power lines) and electronics circuits (like switched-mode power supplies).

But, even if that wasn't the case, there are plenty of naturally occurring sources of EM energy being added to free space including energy from lightning strikes during storms, energy from the sun, and even the residual energy left-over from the Big Bang 13.7 billion years ago (that's the hiss you can hear when you tune between AM radio stations or to an unused channel).

All of these other sources of EM energy in the medium add to the EM energy of any wireless communications signals. Importantly, this changes the instantaneous value of the EM energy of a radio signal, resulting in an instantaneous change of amplitude of the received signal voltage on the output of the antenna.

For a dramatic illustration of the problem for AM, a high energy EM source like a lightning strike would look like a transient (or "spike") superimposed onto the signal at the input to the envelope detector. An example of this is shown in Figure 17 below.

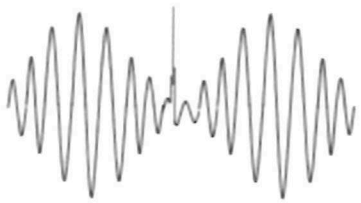


Figure 17 *AM signal with a transient added in the medium caused by a lightning strike*

As the envelope detector is designed to track the received AM signal's envelope, this transient would change the recovered message adding a transient to that too (though possibly not as quickly changing). This is shown in Figure 18 below.

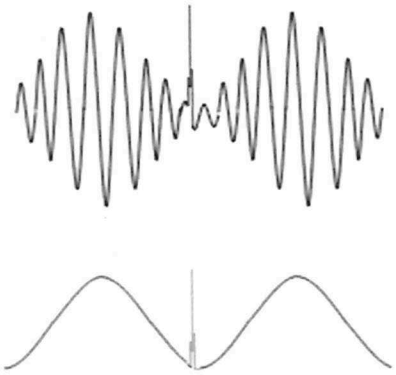


Figure 18

This transient would be heard as a pop that would interfere with the sound of the speech or music out of the radio.

Although this example has used the sudden and large addition of EM energy dumped into the medium caused by a lightning strike, all sources of EM energy in the medium add to the wireless AM signal by some amount. If sufficiently large at the receiver, these other sources of EM energy will also introduce noise interference and may be heard as a hum or buzzing (low frequencies), crackles, whining (high frequencies) or hissing (white noise).

All wireless communications are affected by the addition of EM energy in the medium - Recall that this issue is mentioned when you were introduced to the diagram of a communications system. Some communications systems are more affected than others. Other modulation schemes are inherently better at rejecting this type of introduced noise and FM is an excellent example of this.

Noise and product detection of AM signals – Not tested
Although a product detector can be used to recover the message from DSBFC AM signals, it has exactly the same problem with superimposed noise as the envelope detector and so is not advantageous to use in this regard.

Review questions

Answer these questions to check your understanding of what you have learnt from this chapter. Doing this will help you prepare for the knowledge assessment.

1. Briefly explain the principle of amplitude modulation.
2. What type of non-linear device can generate AM signals?
3. Using a non-linear device to generate AM signals typically produces more sinewaves than are strictly needed for an AM signal. What type of circuit is used to remove them?
4. Sketch a sinusoidal carrier amplitude modulated by a sinusoidal message in the time domain (ie as you would see it on the display of an oscilloscope). Assume a modulation index above 0 (or 0%) and below 1 (or 100%).
5. What does the abbreviation DSBFC stand for?
6. A particular DSBFC system has a 2MHz carrier modulated by a 20kHz message signal:
    a. Calculate the upper and lower sideband frequencies.
    b. Sketch the Am signal in the frequency domain.
    c. Calculate the AM signal's bandwidth.
    d. State whether the AM signal's bandwidth would be different for a complex message having a maximum frequency of 20kHz **and** explain your answer.
7. Determine the modulation index for each of the three AM waveforms shown in Figure 1 below.

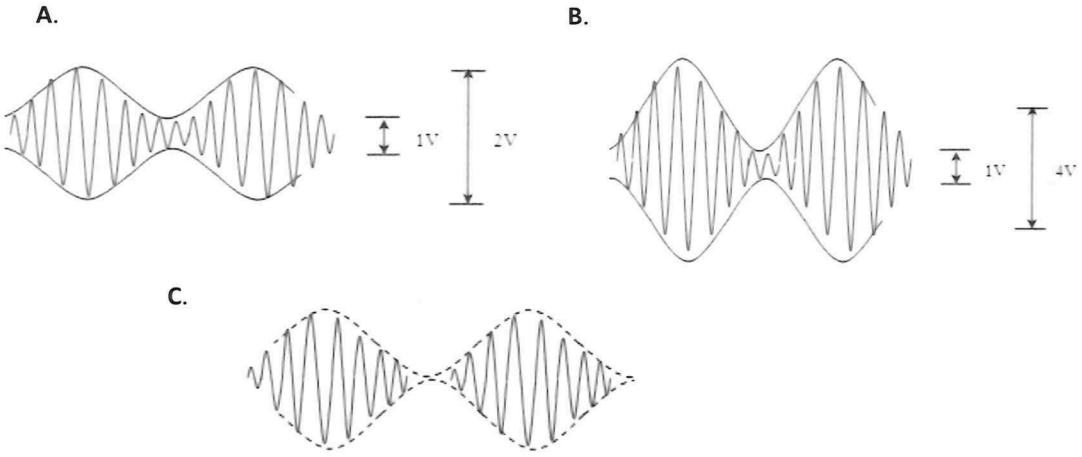


Figure 1

8. Why is it desirable for a transmitter to use modulation levels approaching 100%?
9. What is meant by the term *overmodulation*?
10. Why is overmodulation undesirable?
11. What does the abbreviation SSBSC stand for?
12. State one advantage and one disadvantage of DSBFC over SSBSC.
13. Calculate the total power when a 20kW carrier is DSBFC modulated to 88%.
14. Calculate the total sideband power of the AM signal described in the previous question.
15. Calculate the carrier power for an AM transmission that has a total power of 5kW and a modulation index of 95%.
16. Sketch a simple envelope detector and describe how it works.
17. State two advantages of the using envelope detector to demodulate AM signals.
18. Why can't the envelope detector be used to detect SSBSC?

Student notes

Section 5 Frequency modulation

This section aims to develop you understanding of the principles of frequency modulation (FM), transmission, and how to perform FM modulation index and bandwidth calculations.

At the end of this section you should be able to:

a) Explain the general principle of frequency and phase modulation.

b) Name the type of circuit that can be used to directly generate an FM signal.

c) Define the term *carrier deviation* and state the carrier deviation used for Australian FM broadcast radio.

d) Define modulation index as it applies to FM signals.

e) Calculate the modulation index for an FM system given the carrier deviation and the frequency of the message.

f) State two applications for a table of Bessel Functions with regard to FM signals.

g) Explain the difference between narrowband FM (NBFM) and wideband FM (WBFM) in terms of modulation index and give a widely used application for each.

h) Calculate the approximate bandwidth of an NBFM transmission given the frequency of the message (or the highest frequency in the message).

i) Use Carson's Rule to calculate the approximate bandwidth of a wideband FM transmission given the maximum carrier deviation and frequency of the message (or the highest frequency in the message).

j) State the relationship between the message signal's amplitude the resulting FM signal's modulation index and bandwidth.

k) Explain that the carrier and sideband powers add up to the unmodulated carrier power.

l) State the key difference between FM and FSK.

m) Use an oscilloscope to observe an FM signal (modulated by a sinusoidal message) and determine its modulation index by measurement.

Angle modulation

Previous chapters have explained that modulation involves using a high frequency signal (typically a sinusoid) to "carry" a lower frequency message signal and you have already seen how AM does that. Angle modulation is a similar but different approach to modulation that includes: frequency modulation (FM); and phase modulation (PM).

With FM, the frequency of the carrier is varied by the message signal's amplitude. With PM, the phase of the carrier is varied by the message signal's amplitude. FM and PM are interrelated because phase changes appear as frequency changes and vice versa. (Note: If either the instantaneous frequency or phase is changed, the angle of the instant changes too, hence both are types of angle modulation.)

Frequency modulation

For FM, as the amplitude of the message signal increases, the frequency of the carrier increases proportionally. And, as the amplitude of the message signal decreases, the frequency of the carrier decreases proportionally. This is shown in Figure 1. Notice that the amplitude of the FM signal's carrier doesn't change.

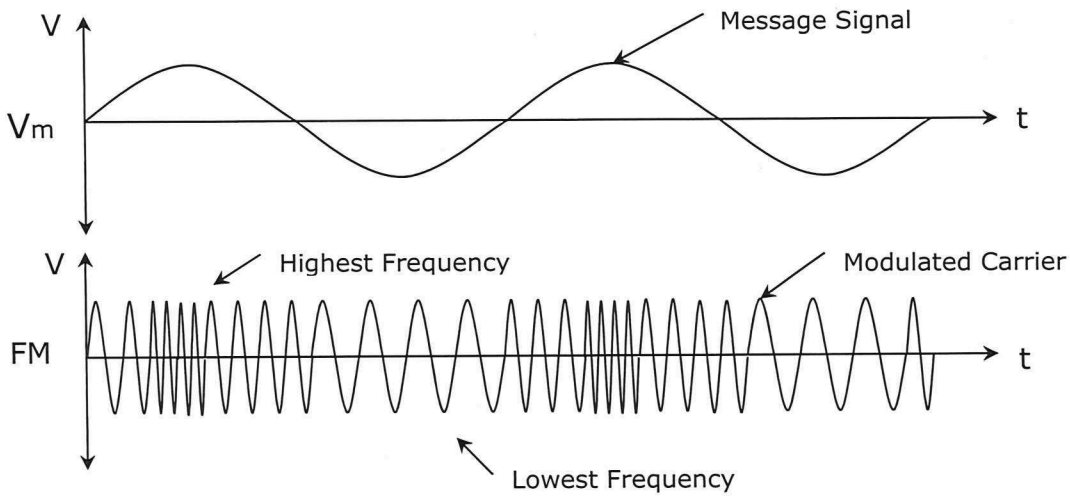


Figure 1 *The message and carrier signals time-coincident with each other*

FM generation

A direct method of generating an FM signal involves using a voltage controlled oscillator (VCO) which you have learnt about before. Recall that a VCO is a sinusoidal oscillator that outputs at set frequency. However, the frequency of the output signal can be varied up and down in response to an input voltage - This is the very essence of FM.

A VCO circuit that can be used to implement FM is the Clapp Oscillator VCO. This circuit is shown in Figure 2 below and its operation would have been discussed when you learnt about Oscillators (EC0064). (See the teacher for a copy of the explanation of this circuit if you have not yet studied EC0064.)

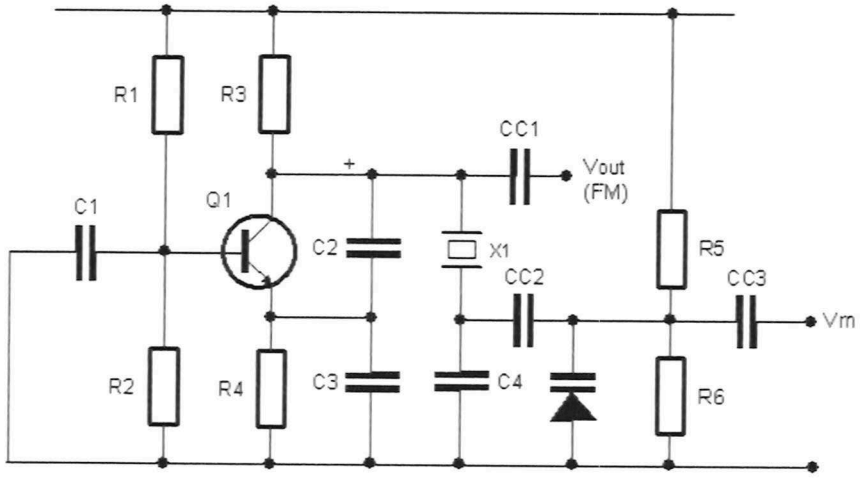


Figure 2 *Direct FM generation using a VCO*

Recall that the Clapp Oscillator is a modification of the Colpitts Oscillator. The crystal in a crystal-controlled Colpitts Oscillator sets the frequency of oscillation but with a crystal-controlled Clapp Oscillator, this frequency is offset by the introduction of a capacitor in series with it (*C4*).

A message signal at the input *Vm* varies the DC voltage across the varicap which varies its capacitance proportionally. As the varicap is connected in parallel with *C4*, changes in the varicap's capacitance causes proportional changes in the total capacitance that is in series with *X1*. This in turn cause proportional changes in the frequency at which the circuit oscillates.

Note 1: This is just one method of implementing a VCO. There are dozens of other circuits that can be used to implement a VCO for direct FM generation.

Note 2: FM signals can be generated indirectly using phase modulation techniques with the carrier modulated by the integral of the message.

Carrier deviation

The term "carrier deviation" is the name for the **amount** by which the carrier frequency (f_C) increases and decreases either side of the unmodulated carrier frequency (ie the "rest frequency" which is also often called the "centre frequency"),

Two symbols are widely used to represent carrier deviation in equations and they are δ or often Δf (your author prefers the latter).

There is no universal mathematical connection between the message signal's amplitude and the amount of carrier deviation. It's governed entirely by the design of the FM generator and its "sensitivity" or "proportionality constant" (expressed in Hz/V).

To give an example, suppose a 100kHz VCO with a sensitivity of 1kHz/V is used to directly generate an FM signal. And suppose the message connected to the VCO is a 5kHz 2Vpp sinewave. The VCO's deviation from its rest frequency will be ±1kHz because the message signal's positive peak is +1V producing up to +1kHz of deviation and its negative peak is -1V producing -1kHz of deviation.

Put another way, the VCO's output frequency will vary from 100kHz up to 101kHz as the message voltage goes from 0V to +1V, then back down to 100kHz as the message returns to 0V, then continue down to 99kHz as the message reaches its negative peak, then return to 100kHz when the message returns to 0V. (Refer to Figure 1 for a visual representation of this frequency change.)

A similar but different attribute of an FM signal is its **rate** of deviation. The frequency of the message sets the rate of the carrier's deviation. For example, a 5kHz modulating signal will cause the carrier to swing above and below its centre frequency five thousand times per second.

Importantly the carrier deviation that is permissible on the channel is determined by ACMA; license holders must not use a carrier deviation that is greater than permitted for the license. For example, Australian broadcast FM radio stations are licensed for maximum carrier deviations of ±75kHz. (However, this **does not mean** that the channel bandwidth of Australian broadcast FM is 150kHz. It's more complicated than that as will be explained next.)

Modulation index

The modulation index for FM (m_f or sometimes β) is a measure of the extent to which the carrier is varied by the message signal and can be found as follows:

$$m_f = \frac{\delta}{f_i} \qquad \text{or} \qquad m_f = \frac{\Delta f}{f_m}$$

Where:

$\delta = \Delta f =$ Carrier deviation and

$f_i = f_m =$ Message frequency

For example, the modulation index for an FM system with a maximum deviation of ±50kHz and a 20kHz message signal is:

$$m_f = \frac{\Delta \mathrm{f}}{\mathrm{f_m}}$$

$$m_f = \frac{50kHz}{20kHz}$$

$$m_f = 2.5$$

Importantly, for AM, the modulation index is used to determine the distribution of the modulated AM signal's total power between the carrier and the sidebands. However, for FM, the modulation index is used for something completely different: Bandwidth calculations. This is discussed next.

Bandwidth requirements for FM transmission

As already explained, when a carrier is frequency modulated by a single tone, the carrier is caused to swing above and below its rest frequency. (This is true for complex message signals too but it's much simpler to just talk about the message being a sinewave otherwise it all gets too complicated for an introduction to the topic.)

Importantly, when the carrier deviates from its rest frequency, a theoretically infinite number of sidebands are produced at frequencies above and below the carrier. (This is difficult to prove without some seriously heavy duty maths so you're going to have to take this fact on trust!)

Fortunately, most of these sidebands are insignificant otherwise the required bandwidth would also be infinite (which means multiple users would be unable to share the spectrum). That said, some sidebands are significant and the modulation index is used to identify them and, in turn, determine the FM signal's bandwidth.

Unlike AM, the calculations to determine the significant sidebands for a given modulation index of an FM signal are very complex. Instead, the table of *Bessel Functions* can be used used. (This is true even for engineers and if you look-up "Bessel function" in Wikipedia, you'll see why!)

The table of *Bessel Functions* lists a set of modulation indexes and specifies the size of the carrier - labelled J_0 - and pairs of sinewaves that straddle the carrier labelled J_1 (for the pair closest to the carrier on either side), J_2 (for the next pair of sinewaves either side) and so on.

An excerpt of the table of Bessel Functions is shown below.

	CARRIER	Sidebands											
m_f	J_0	J_1	J_2	J_3	J_4	J_5	J_6	J_7	J_8	J_9	J_{10}	J_{11}	J_{12}
0.00	1.00	—	—	—	—	—	—	—	—	—	—	—	—
0.25	0.98	0.12	—	—	—	—	—	—	—	—	—	—	—
0.5	0.94	0.24	0.03	—	—	—	—	—	—	—	—	—	—
1.0	0.77	0.44	0.11	0.02	—	—	—	—	—	—	—	—	—
1.5	0.51	0.56	0.23	0.06	0.01	—	—	—	—	—	—	—	—
2.0	0.22	0.58	0.35	0.13	0.03	—	—	—	—	—	—	—	—
2.5	−0.05	0.50	0.45	0.22	0.07	0.02	—	—	—	—	—	—	—
3.0	−0.26	0.34	0.49	0.31	0.13	0.04	0.01	—	—	—	—	—	—
4.0	−0.40	−0.07	0.36	0.43	0.28	0.13	0.05	0.02	—	—	—	—	—
5.0	−0.18	−0.33	0.05	0.36	0.39	0.26	0.13	0.05	0.02	—	—	—	—
6.0	0.15	−0.28	−0.24	0.11	0.36	0.36	0.25	0.13	0.06	0.02	—	—	—
7.0	0.30	0.00	−0.30	−0.17	0.16	0.35	0.34	0.23	0.13	0.06	0.02	—	—
8.0	0.17	0.23	−0.11	−0.29	−0.10	0.19	0.34	0.32	0.22	0.13	0.06	0.03	—
9.0	−0.09	0.24	0.14	−0.18	−0.27	−0.06	0.20	0.33	0.30	0.21	0.12	0.06	0.03
10.0	−0.25	0.04	0.25	0.06	−0.22	−0.23	−0.01	0.22	0.31	0.29	0.20	0.12	0.06
12.0	0.05	−0.22	−0.08	0.20	0.18	−0.07	−0.24	−0.17	0.05	0.23	0.30	0.27	0.20
15.0	−0.01	0.21	0.04	−0.19	−0.12	0.13	0.21	0.03	−0.17	−0.22	−0.09	0.10	0.24

There are some important/and or useful points that need to be made about this table.

1) The figures to the right of the modulation index (m_f) represent the amplitude of the carrier and sidebands as a proportion of the unmodulated carrier (hence, why the carrier is 1 when the modulation index is 0).

2) If you use the simple sum to combine the figures along a line you'll find you don't get "1". For example, if you added the figures for $m_f = 0.25$, you see that doing so would result in 1.1 (ie 0.98 + 0.12). This is because the sinewaves are not all at the same frequency and so their phases must therefore all be different and changing continuously too.

3) The appropriate mathematical sum of the carrier and sideband amplitudes always equals one. This fact tells you an important thing about FM – The power in an FM signal is constant regardless of the modulation index. (More on this later.)

4) The greater the depth of modulation (ie the bigger the modulation index) the more significant sidebands there are in the FM signal. This in turn means that, the greater the depth of modulation, the wider the FM signal's bandwidth.

5) In this table, "significant" sidebands are all sidebands with a minimum amplitude of 0.01 (ie 1%) or more of the unmodulated carrier's amplitude. It is an engineering decision about whether to use this table of Bessel Functions, or a different one showing sidebands with a lower or higher minimum amplitude.

6) The minus sign indicates 180° of phase shift.

7) The side frequencies (J_1, J_2, etc) are spaced equally at the frequency of the message. For example, if the message is 5kHz, J_1 represents a pair of sinewaves that are ±5kHz about the carrier, J_2 represents a pair of sinewaves that are ±10kHz about the carrier, and so on.

8) There are modulation indexes where the carrier component (J_0) is insignificant. To illustrate this, look down the table for J_0 at different modulation indexes. For example, for $m_f = 0.25$ the carrier component makes up only 1% of the total power in the FM signal.

It's useful here to talk to some of the points on the previous page using frequency domain graphs. To this end, suppose a 1MHz sinusoidal carrier is frequency modulated by a 20kHz sinusoidal message. For context, Figure 6 below shows the spectral composition of the unmodulated carrier – It's just a sinewave at 1MHz (J_0) and is obviously 100% of its amplitude.

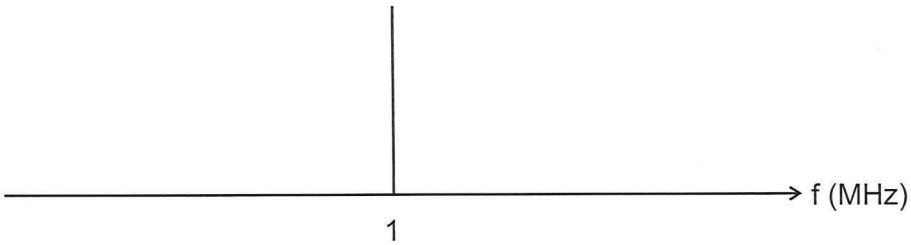


Figure 6 *An unmodulated 1MHz carrier in the frequency domain*

Now suppose this carrier is frequency modulated by the 20kHz sinusoid to a depth of modulation of $m_f = 0.25$. Figure 7 below shows the spectral composition of the FM signal. As you can see, it consists of the carrier (J_0) which is 2% smaller than the unmodulated carrier's amplitude and just one pair of sidebands (the J_1 pair). Interestingly, the spectral composition of this FM signal looks just like the spectral composition of a DSBFC AM signal and so the bandwidth of an AM signal and an FM signal for $m_f = 0.25$ are the same (40kHz in this example).

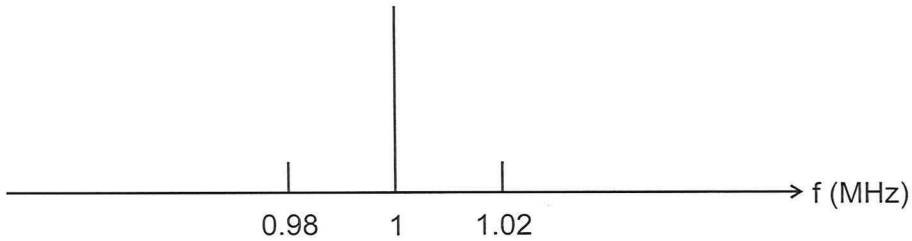


Figure 7 *The spectral composition of the FM signal for $m_f = 0.25$*

Now suppose the depth of modulation is increased to $m_f = 0.5$. Figure 9 below shows the new spectral composition of the FM signal. As you can see, it consists of the carrier (J_0), the J_1 sidebands, and the J_2 sidebands. The bandwidth of this FM signal (80kHz) is double that for $m_f = 0.25$. Notice that J_0 is a little bit smaller again in Figure 8 and the J_1 sidebands are bigger. This reflects their sizes as specified in the table of Bessel Functions.

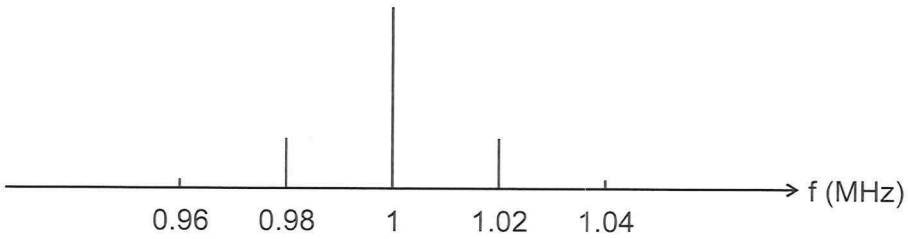


Figure 8 *The spectral composition of the FM signal for $m_f = 0.5$*

Now suppose the depth of modulation is increased to $m_f = 1$. Figure 9 below shows the new spectral composition of the FM signal and the components' relative sizes. Notice that the FM signal's bandwidth is now 120kHz, the carrier is smaller still and the J_1 sidebands are much bigger. Again, this reflects the figures in the table of Bessel Functions.

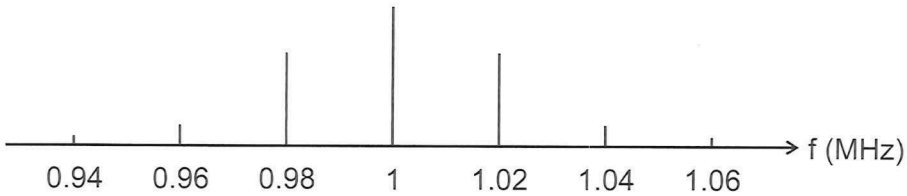


Figure 9 *The spectral composition of the FM signal for $m_f = 1$*

FM bandwidth calculations

It would be more convenient to determine the bandwidth of an FM signal without having to refer to the table of Bessel Functions. As already mentioned, for an accurate figure for any value of m_f, this involves some heavy-duty maths.

Fortunately, there are two approximations for bandwidth: one for Narrow Band FM (NBFM) which is used for UHF two-way FM radio communications (which typically just involves speech); and another one for Wideband FM (WBFM) which is used for broadcast radio (speech and music).

By definition, NBFM is any FM communications system with a modulation index equal to or less than 0.5 (that is, $m_f \leq 0.5$). For NBFM, the signal's bandwidth can be approximated to:

$$\boxed{BW = 2 \times fm}$$

Where: *fm* is the message frequency

As noted above, this equation is the same as the bandwidth calculation for DSBFC AM.

By definition, WBFM is any FM communications system with a modulation index equal to or greater than 5 (that is, $m_f \geq 5$). For WBFM, the signal's bandwidth can be approximated using an equation called Carson's Rule:

$$\boxed{BW = 2(f_m + \Delta f)}$$

For example, let's consider broadcast FM radio. As you know, this system uses a deviation of ±75kHz with a maximum allowable message frequency of 15kHz (so $m_f = 5$). According to Carson's Rule, the bandwidth of broadcast FM radio signals is approximately:

$$BW = 2(f_m + \Delta f)$$

$$BW = 2(15kHz + 75kHz)$$

$$BW = 180kHz$$

As an aside, Carson's Rule specifies a bandwidth that includes the power in the carrier and sufficient sidebands to equal 98% of the transmitted power. This is an engineering compromise – using a wider system bandwidth for FM would result in capturing only 2% more of the transmitted power but permit fewer channels in the band allocation.

A question you're possibly asking yourself now is, why bother with WBFM if it produces a modulation signal with such a wide bandwidth? Wouldn't it reduce the number of users in the band?

You're right. For the range of frequencies in the FM band, you could cram many more NBFM channels in than WBFM channels.

However, as you will see, FM is relatively immune to noise in the channel (and noise introduced at the receiver too). Importantly, using WBFM further improves this noise performance advantage of FM for the higher frequencies in complex message signals like music.

Power in FM signals

As already mentioned, when the sum of the carrier and the sidebands is determined correctly, the total always equals one. This means that the power in an FM signal is constant regardless of the modulation index.

This fact leads to another really important advantage of FM over AM. The power in the carrier isn't wasted like it is in AM. So, for all other things being equal (eg antenna length, antenna polarisation, antenna height above the ground, wave propagation method use, etc, etc) a 150W FM transmission will carry much further than a 150W AM transmission. All of the 150W in the FM signal carries the message whereas, only 50W (at best if the modulation index is 1!) of the 150W AM signal carries the message.

Frequency shift keying

As you know, the message signal in communications systems can be digital data. When data is used as the message in frequency modulation, the system becomes known as *frequency shift keying* (FSK).

Figure 10 below shows a simple squarewave being used to represent a message that is data. Time-coincident with the message is the frequency modulated carrier (the FSK signal).

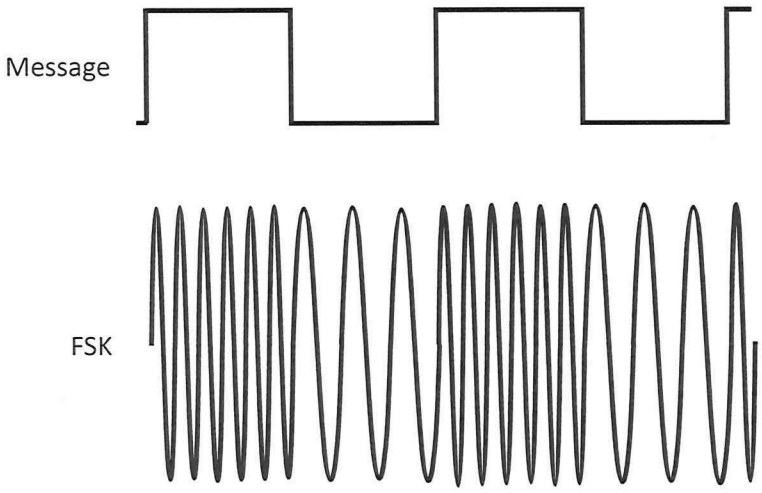


Figure 10 *A squarewave message time-coincident with the carrier that it is being used to frequency modulate*

It hasn't been discussed yet but FM communication systems have a significantly better noise performance than AM communication systems. This means that FSK must have a significantly better noise performance than ASK and so it might generally be preferred. That said, unless NBFM is used, FSK with high data rates needs a wide channel bandwidth to accommodate all of its significant sidebands.

In truth, ASK and FSK are mostly only used in niche applications these days (like telemetry) as phase shift keying (PSK) and its variations are preferred. PSK has the same kind of noise performance as FSK but has better bandwidth efficiency. That said, apparently FSK is still being used in legacy telephony systems (that is, systems that were developed many years ago that are still in use today), like the Caller ID system on landline handsets.

Review questions

Answer these questions to check your understanding of what you have learnt from this chapter. Doing this will help you prepare for the knowledge assessment.

1. Describe the relationship between the message signal and modulated carrier in an FM system.
2. What type of circuit can directly implement FM?
3. Describe the term *carrier deviation* as applied to FM.
4. What is the *rate* of deviation of an FM signal dependent on?
5. What is the relationship between the message signal amplitude and the amount of deviation of an FM carrier?
6. Determine the two frequencies between which a 100MHz carrier swings if the deviation is 50kHz.
7. What amount of carrier deviation used in broadcast FM?
8. Determine the modulation index (m_f) for an FM system having a carrier deviation of 15kHz caused by a message signal of 5kHz.
9. What does a table of Bessel Functions show?
10. What can the table of Bessel Functions be used for?
11. What happens to the sidebands as the modulation index increases? (Consult a table of Bessel Functions for the answer.)
12. Calculate the approximate bandwidth of a narrowband FM signal having a deviation of ±1.5kHz for a maximum frequency of 3kHz.
13. Calculate the modulation index of the FM signal for the previous question.
14. Why does the FM signal qualify as NBFM?
15. Calculate the approximate bandwidth of a WBFM FM signal having a carrier deviation of ±75kHz for a maximum message frequency of 12kHz.
16. Calculate the modulation index of the FM signal for the previous question.
17. Why does the FM signal qualify as WBFM?
18. Explain why power in an FM signal is constant.
19. How does FSK differ to FM?

Section 6 FM demodulation

This section aims to develop your understanding of the principles of FM demodulation.

At the end of this section you should be able to:

a) State the other name that is commonly used for many FM demodulator circuits.

b) Describe the key performance characteristic of an ideal FM demodulator.

c) Explain why FM has an inherently better noise performance compared to AM.

d) Give one of the two reasons for using a limiter stage in an FM receiver.

e) State that noise can still affect FM communications despite its better noise performance compared to AM.

f) State which components in a complex message are most affected by noise in FM communications.

g) Briefly describe what pre-emphasis does to the message signal before it is used to modulate the carrier in FM communications.

h) Briefly explain what de-emphasis does to the recovered message signal in FM communications.

i) Name and/or draw the circuit that can be used to implement de-emphasis and state the time constant used for the circuit in the receivers of Australian broadcast FM radio.

j) Name one of the discriminator circuits used to demodulate FM signals in older FM radio communication systems.

k) Draw the block diagram of the phase-locked loop (PLL) and explain its basic operation as an FM demodulator.

l) State the main advantage of the PLL over the FM discriminators for demodulating FM signals

The principle of FM demodulation

An FM demodulator, often called a *discriminator*, recovers the message signal from the FM carrier. The discriminator does this by converting any deviation in carrier frequency to a proportional voltage. This is the reverse of FM modulation.

Ideally, the discriminator has a linear response between the FM signal's deviation limits. The discriminator's performance beyond those limits is irrelevant. An example is shown in Figure 1 below for a discriminator that could be used to recover the message from Australian broadcast FM radio transmissions.

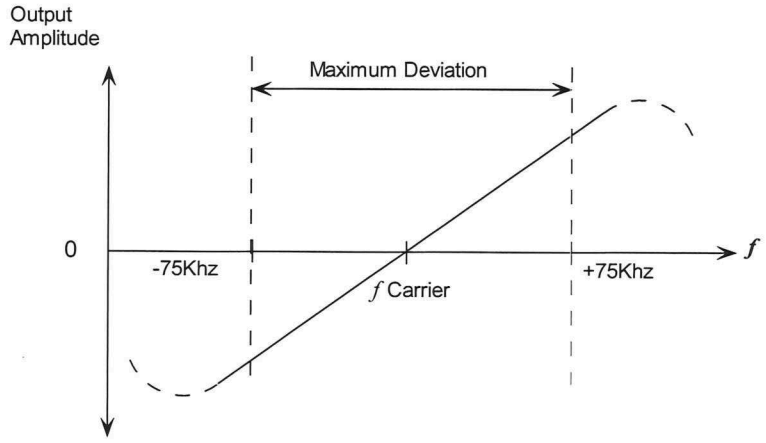


Figure 1 *The response of an ideal FM discriminator*

It can be seen from Figure 1 that, when the carrier is at its centre frequency, the discriminator's output is zero. As the carrier swings above the centre frequency, the discriminator's output increases amplitude in the positive direction. As the carrier swings below its centre frequency, the discriminator's output increases in amplitude but in the negative direction. This means that deviations in carrier frequency are converted to proportional amplitude variations. And, as the deviations in the carrier are caused by the amplitude variations of the original message, the output of the discriminator must be a copy of the original message.

Several different FM demodulator circuits have been commercially used over the years and these are discussed later in this chapter.

FM and noise

Generally speaking, FM communications has superior noise performance characteristics compared to AM communications. That is, the same amount of noise in the medium, or introduced by the circuits at the input of the receiver, is much less noticeable for FM communications systems than for AM communications systems.

One reason for this is relatively easy to explain. Recall that noise added to signals causes changes in the signal's amplitude. Unfortunately, for DSBFC AM, the envelope detector is designed expressly to detect amplitude variations so they detect this noise too. Obviously then, any noise frequencies in the audio range can be heard. Noise in ASK signals has the potential to corrupt data bits by causing the voltage to be different enough to look like the opposite logic level at the moment that the data bit is read.

However, for FM, it's the carrier's frequency that is varied in response to the message signal voltage. That being the case, the demodulation circuits must be designed to look for frequency changes and can ignore amplitude variations (ideally). In the process, any amplitude variations caused by noise are ignored too.

That said, FM is still somewhat affected by amplitude variations caused by noise for reasons discussed shortly.

Limiter stage

A number of FM demodulation circuits (discriminators) have been used over the years. Some of them respond not just to frequency changes (desirable) but to amplitude changes too (undesirable). On account of this, these demodulators are not as immune to noise as they could be.

However, a neat solution for the affected demodulators involves the inclusion in the receiver of a circuit called a *limiter* (situated in the receiver prior to the FM demodulator). The limiter literally chops-off the tops and bottoms of the FM signal. In so doing the limiter chops-off any amplitude variations due to noise and so there are no amplitude variations for these FM demodulators to respond to. An example of a limiter block is shown in Figure 2 below.

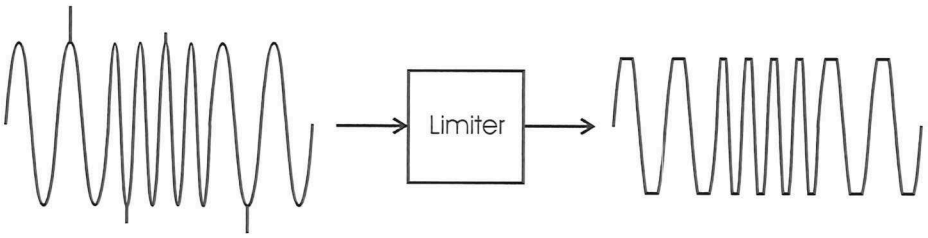


Figure 2 *A limiter chops-off the top and bottom of the received FM signal to remove amplitude variations due to noise*

A simple method of implementing limiting involves an amplifier with a gain so large that the received FM signal overdrives the amplifier causing clipping.

A secondary benefit of limiting is that it virtually eliminates the need for AGC in FM receivers because the signal from the limiter to the demodulator is always the same size, regardless of how small the received signal was. So, a limiter may be included in an FM receiver design even when the discriminator is not one that is affected by amplitude variations.

Pre-emphasis and de-emphasis

Although FM receivers have inherent noise immunity, noise still can affect FM signals. This can be understood using math that can be a little challenging (for your author, at least). So phasor diagrams can be used instead.

Recall that a phasor diagram represents a sinewave as a rotating line. The instantaneous position of the line's tip represents the instantaneous voltage (or current) of the sinewave. An example of a phasor is shown in Figure 3 below. Its position points to the sinewave's instantaneous positive peak voltage.

(If you're a little rusty on phasors, you may want to read the notes in the linked website below before continuing (but only read down to the paragraph that starts "Mathematically we can add the two voltages together by"– don't tackle the math that follows that unless you have learnt about imaginary numbers (also called complex numbers and j-notation): https://www.electronics-tutorials.ws/accircuits/phasors.html)

If the sinewave that the phasor in Figure 3 represents is 1MHz then the phasor completes one million revolutions per second. This obviously can't be shown using ink on a page so you have to picture the line spinning around (and doing so very quickly if the sinewave is 1MHz!).

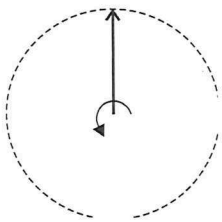


Figure 3

Let's assume that this 1MHz sinewave is the carrier in an FM system modulated by a sinusoidal message with a magnitude that causes a deviation of ±50kHz. This means that the rotational speed of this phasor is speeds up when the carrier deviates above its rest frequency and slows down when it deviates below its rest frequency. Again this is something you have to imagine.

Now, let's assume that noise is added to this 1MHz carrier. Most noise consists of many sinewaves but, for ease of explanation, let's also assume that the noise is a single sinewave and has a frequency of 1.1MHz. As the noise is added to the carrier, its phasor sits on top of the carrier's phasor and is rotating anti-clockwise as well, one million one hundred thousand times per second (in this example). As the two phasors rotate at different speeds, they seldom point in the same direction as each other. That being the case, freezing the phasors at different times as shown in Figure 4 below (always when the carrier is at its peak for ease of comparison) would likely show the noise phasor in different locations each time.

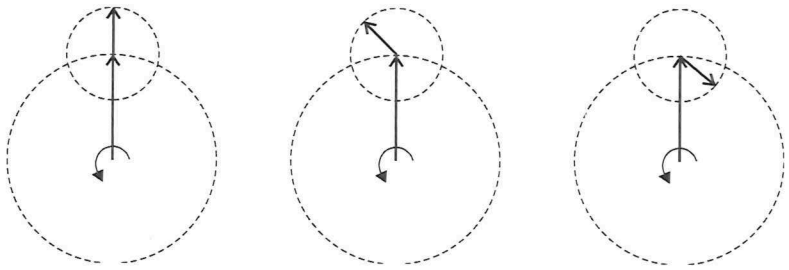


Figure 4

As the carrier's instantaneous value is the vectorial sum of the two phasors, the resulting carrier for the three instants in Figure 4 are shown in Figure 5 below.

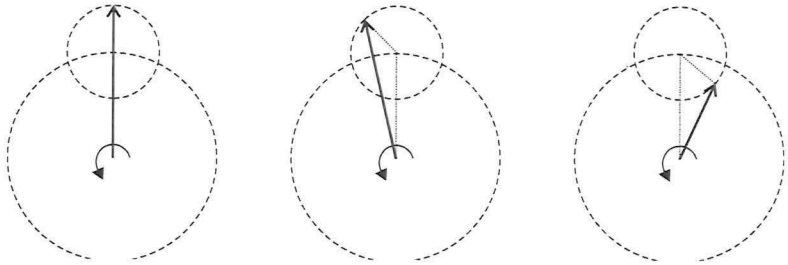


Figure 5

Notice that the noisy carrier's phasor is sometimes longer than the carrier's phasor without noise and sometimes shorter. (The noisy carrier is longer in the left-side and middle diagrams and shorter in the right-side diagram.) This represents the changes in carrier amplitude that noise introduces which is either ignored by the FM demodulator or is chopped-off at the top and bottom by a limiter. So, no problems there.

However, do you notice something else? The resulting carrier's phase angle in the left-side diagram in Figure 5 hasn't changed but it has changed for the other two diagrams. For the middle diagram, the carrier's phase angle is ahead of where it should have been without the noise. For the right-side diagram in Figure 5, the carrier's phase is behind where it should have been without the noise.

This is significant because, during the FM signal's phase transition (due to the effect of the noise sinewave), the carrier looks like it's changing frequency too because, while the phase changes, the time changes too. This in turn makes the carrier's deviation look bigger or smaller than what it would have been without the noise.

As discriminators are designed to look for frequency deviations, any errors in frequency deviation (due to the noise-induced phase shifts) produce corresponding errors in the instantaneous recovered message voltage compared to the instantaneous original message voltage. Hence noise still ends up being added to the recovered message (though, it's not nearly as bad as the effect of the amplitude variations that noise causes in an AM system).

This example has been explained using a single noise sinewave but, in reality, there would be many noise sinewaves which would be heard as hiss out of the receiver if the noise signal is large enough.

In summary, noise ultimately causes unwanted frequency deviations as well as amplitude deviations. And as discriminators are designed to respond to frequency deviations, they demodulate this effect of noise and so it can still be heard even when a limiter is used.

The problem is more complicated than that because, as the breakout box on the next page explains, the noise affects higher message frequencies more than lower frequencies. That is, suppose you were transmitting a 500Hz tone using FM and you could hear some noise. The exact same level of noise in the system would sound louder if the tone being transmitted was 5kHz and louder again if the tone was 15kHz.

This obviously has implications for transmitting complex message signals with a wide bandwidth like music.

And, complicating the matter a little more, the frequency components in music generally get smaller the higher their frequency. This means that the carrier frequency deviation they produce is smaller in the first instance than it is for lower frequencies and this makes the deviation error for the higher frequencies relatively even larger.

Why noise in FM affects higher frequencies more – Not tested

As already mentioned, FM inherently cancels the amplitude variations caused by noise but does respond to the frequency deviations that noise causes because it causes spontaneous phase shifts and while the phase changes, the frequency appears to change too.

That said, even with regard to noise-induced frequency deviations, WBFM is an inherently good performing system. The noise added to the modulated carrier (either in the medium or at the receiver) would have to be substantial to be heard via this phase-shifting mechanism. In fact, WBFM has inherent cancelling of noise caused by this mechanism – A SNR of 0.5 at the input to the receiver results in an SNR of 0.1 in the recovered message.

An understanding of this induced-noise phase error leads to a more important issue which was referred to on the previous page.

Suppose the phase error caused by the noise is +10°. If the message frequency was a single sinewave at 500Hz and was an instantaneous voltage that was supposed to produce a carrier deviation of say 50kHz, the math (not shown here) predicts that the error caused by the noise would be 87Hz. That is, the carrier's instantaneous deviation would be 50,087Hz instead of 50kHz (a frequency error of 0.174%).

But if the message frequency was a single sinewave at 5kHz and was an instantaneous voltage that was supposed to produce a deviation of 50kHz, the math predicts that the error caused by the noise would be 870Hz. That is, the carrier's instantaneous deviation would be 50,870Hz instead of 50kHz (a frequency error of 1.74%).

From these figures we can see that the deviation error is directly proportional to the message frequency. The higher the message frequency (for a given instantaneous message voltage), the greater the frequency deviation error at the receiver, and hence the louder the noise.

The solution to this problem is a signal conditioning technique at the transmitter called *pre-emphasis*. Before the message signal is used to modulate the carrier, the complex speech/music signal is modified so that the amplitude of the higher frequency components is made bigger than normal (a bit like you do when you turn up the treble on an amplifier at home, but done at the transmitter).

This means that the higher frequencies cause much greater carrier deviations than they would do normally. But, as the deviation error at the receiver caused that is caused by the noise is (in part) a function of the message frequency, the absolute value of the deviation error remains the same and so is proportionally a smaller amount of the total deviation.

To explain that last paragraph using figures – Not tested
Suppose the amplitude of our 5kHz sinewave discussed in the break-out box on the previous page is increased by, say, 50%. The carrier deviation would increase by 50% too, to 75kHz. As the carrier deviation error introduced by the noise would be unchanged (remember, it's a function of the phase error introduced by the noise and the message's frequency), a change of 870Hz out of 75kHz represents a 1.16% error instead of 1.74% and so the noise would be smaller and hence less audible.

In practise, the amplitude of the higher frequencies of complex messages like music can be increased by much more than 50% (reducing the effect of the noise even more) because, as already mentioned, they're much smaller than the lower frequency components.

Figure 6 below shows where a pre-emphasis stage would sit in an FM transmitter.

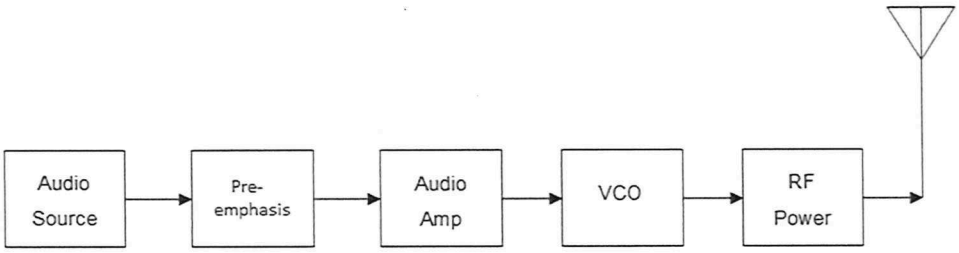


Figure 6 *Block diagram of an FM transmitter with pre-emphasis stage*

For Australian broadcast FM radio, the pre-emphasis stage is designed to increase the amplitude of frequencies above 3.182kHz (as shown in Figure 7) at a rate of +6dB/octave (which improves the FM signal's SNR at frequencies above 3.182kHz at the same rate).

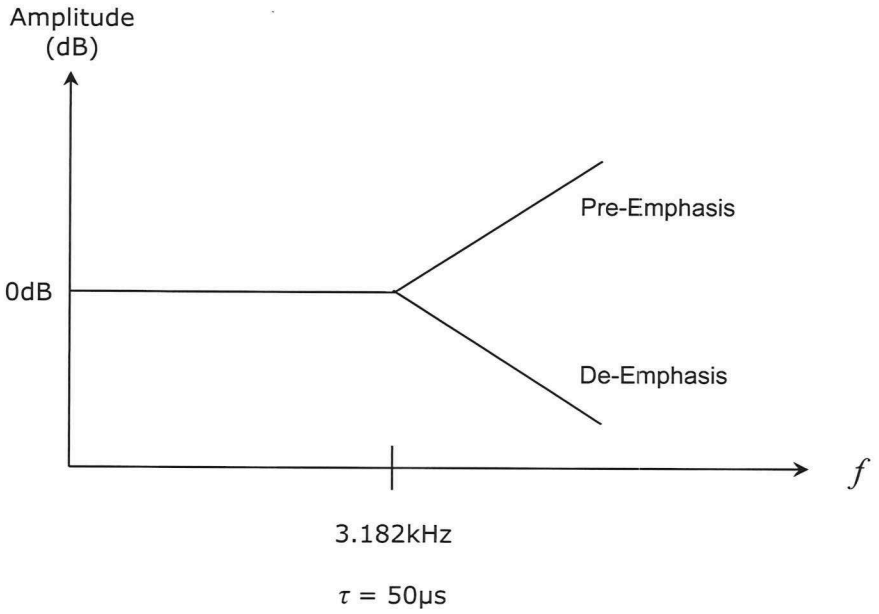


Figure 7 *Pre-emphasis and de-emphasis*

Pre-emphasis must be undone at the receiver in a process is called *de-emphasis* to restore the high frequency components to their original relative levels. Otherwise, the received signal will sound thin and "tinny".

De-emphasis is applied to the recovered message (that is, after the discriminator) by attenuating frequencies above 3.182kHz (as shown in Figure 7) at -6dB/octave.

The de-emphasis circuit is surprisingly simple and is shown in Figure 8 below. It's just a first order RC low-pass filter with a cut-off frequency of 3.182kHz. (Recall that first order filters have a slope of 6dB/octave.)

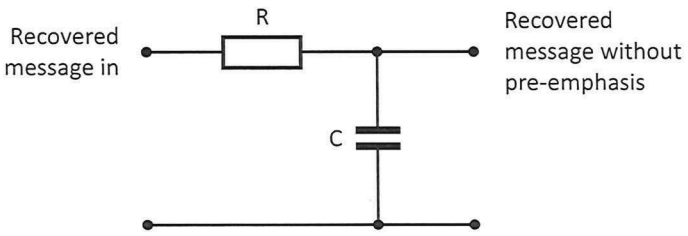


Figure 8 *De-emphasis circuit used in FM receivers*

A curious thing about pre-emphasis and de-emphasis is that the frequency at which it kicks-in is often not quoted as a frequency at all. Instead, it's often quoted as occurring at 50μs.

This figure comes from the equation for finding the cut-off frequency of an RC LPF which is: $fc = 1/(2\times \pi \times R\times C)$. The 50μs is what you get when you just multiply R and C together. Check this for yourself by entering $1/(2\times \pi \times 50\mu s)$ into your calculator - You'll see that you get 3.182kHz.

You may recall that $R \times C$ is the equation for finding the time constant (τ) of a series RC circuit and so this figure is referred to as the pre-emphasis and de-emphasis circuits' time constant.

Not all countries use the same de-emphasis time constant – Not tested
In north America (so, the US and Canada), the time constant used for pre-emphasis and de-emphasis is 75μs and not 50μs as it is here, Asia and in Europe. This means that pre-emphasis is added at the transmitter to frequencies above 2122Hz.

FM discriminators

As previously mentioned, several discriminators have been used over the years. The following notes provide a brief overview. For the final theory test written by the author of this book, you need only know the names of the circuits on this page and the next. A more detailed description is provided for the *Phase-locked Loop's* operation because it is the only FM demodulator that will be assessed in the final theory test (written by the author).

Slope detector

The slope detector uses an LC tank circuit (band-pass filter response) tuned so that the carrier frequency falls in the middle of one of its slopes. Variations in frequency (FM) are therefore converted to variations in amplitude (AM). The message is recovered from the AM signal using a envelope detector.

Slope detectors are one of the easiest discriminators to understand and implement, but they are not particularly linear and so your author doubts that it has been used much for receivers of FM broadcast radio, if at all.

The Foster-Seeley and Ratio detectors

The *Foster-Seeley* and *Ratio* detectors are two FM discriminators that were once the most widely used. Today, their use is diminishing because their implementation cannot be incorporated into an integrated circuit (they both rely on a transformer for their operation).

The Foster-Seeley detector uses a tuned RF transformer to convert frequency changes into amplitude changes. A centre-tapped transformer, tuned to the carrier frequency, is connected to two rectifier diodes in a full-wave bridge rectifier configuration. If the input frequency equals the carrier frequency, the two halves of the tuned transformer circuit produce the same rectified voltage and the output is zero. As the frequency of the carrier changes, the balance between the two halves of the transformer secondary changes and the result is a voltage proportional to the frequency deviation of the carrier.

The Foster-Seeley discriminator is sensitive to both frequency and amplitude variations. Therefore, the FM signal must be supplied to Foster-Seeley discriminator via a limiter stage to remove amplitude variations. The Foster-Seeley detector offers excellent linear response to FM signals having large frequency deviations.

The Ratio detector is a variant of the Foster-Seeley discriminator (where one diode conducts in the opposite direction). The Ratio Detector provides better immunity to unwanted amplitude variations than the Foster-Seeley Detector but its output voltage is significantly smaller.

The quadrature detector

The *quadrature* detector gets its name because the demodulation process uses two signals, with one being 90° out-of-phase (and so, in "quadrature") with the other.

The received FM signal his severely limited and so it resembles a square wave (with a changing frequency) and is fed to one input of an exclusive-OR gate.

This squarewave is also fed to the other input of the exclusive-Or gate via an LCR circuit that filters the harmonics to leave the squarewave's fundamental that is phase-shifted by 90°.

Recall that the XOR gate's output is logic-0 when the inputs are the same and logic-1 when they're different. As the two signals on its input are 90° phase shifted, the result on the XOR gate's output is a series of pulses to logic-1 from logic-0.

Any frequency deviation in the FM carrier causes a corresponding phase deviation in the signal on the LCR circuit's output. And this changes the width of the pulses, decreasing as the phase shift reduces (because the two signals overlap for longer) and increasing as the phase shift increases (because the two signals overlap for less time).

The effect is a pulse width modulated signal that is proportional to the carrier deviation (so in effect, the quadrature detector converts the FM signal to PWM). A low pass filter connected to the output of the XOR gate gives an average value of the PWM signal and is a copy of the original message.

The quadrature detector can be mostly implemented in an integrated circuit and so has been a widely used alternative to the Foster-Seeley and Ratio detectors. That said, it still needs an externally connected tuned circuit.

FM demodulation using the Phase-Locked Loop (PLL)

The *phase-locked loop* (PLL) has become the most widely used means of FM demodulation in recent years as it can be wholly implemented in an integrated circuit. Prior to the advent of the IC, the PLL was not widely used commercially as it is profoundly difficult to implement using discrete components (which made it expensive).

The operation of the PLL is best described using a block diagram as shown in Figure 9.

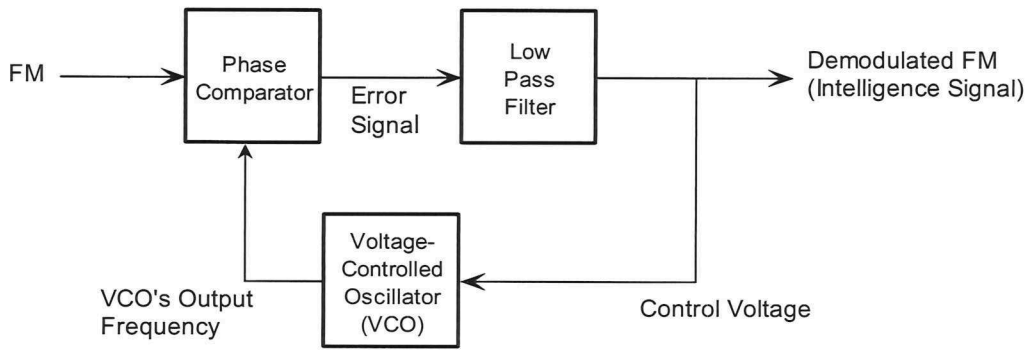


Figure 9 *The block diagram of a phase-locked loop*

The PLL uses a phase comparator to compare the phase of the incoming FM signal with that of a local voltage controlled oscillator (VCO). The phase comparator outputs an error signal that, when low-pass filtered, produces a control voltage for the VCO that is proportional to the difference between its two input signals. (The spectral composition of the error signal is very complex and includes many high frequency components. The LPF is necessary to remove them and prevent them from confusing the VCO.)

To explain how the PLL works, let's start with an unmodulated FM signal on its input. At first, the incoming frequency and VCO frequency are different. This frequency difference appears as a rolling phase difference to the phase comparator and this produces a control voltage that swings up and down. This in turn causes the VCO's output frequency to swing up and down as well.

If the PLL is suitably designed for the task, at some point, the VCO's output frequency will swing past the unmodulated FM carrier's frequency. But, the moment they are the same frequency, the phase difference between the two signals no-longer appears to be changing to the phase comparator and so the control voltage stops changing too. Once that happens, the VCO's output stops changing frequency and sticks at the same frequency as the unmodulated carrier. At that moment, the PLL is said to be in "lock" with the incoming signal.

Once the PLL is in lock, what happens next can probably be guessed by the term.

When the FM carrier is modulated, its frequency deviates up and down producing a phase error again. However, the phase error produces a change in the control voltage that forces the VCO's frequency up and down to match the change in the incoming FM signal. In other words, while-ever the FM signal is changing frequency, the control voltage changes size in the same proportion to force the VCO's output frequency to match it.

Importantly, as FM carrier frequency deviations are a direct result of the original message signal's amplitude, this means that the control signal to the VCO must be a copy of the original message signal.

Review questions

Answer these questions to check your understanding of what you have learnt from this chapter. Doing this will help you prepare for the knowledge assessment.

1. Explain why FM has better noise immunity compared to AM.
2. Explain the purpose of the limiter stage in FM.
3. Explain the purpose of pre-emphasis.
4. Where in an FM communications system is pre-emphasis found?
5. Above what frequency does pre-emphasis take effect in Australia?
6. Explain the purpose of de-emphasis.
7. Where in an FM communications system is de-emphasis found?
8. What circuit is used to implement de-emphasis?
9. What does it mean when it is said that the pre-emphasis and de-emphasis in Australian FM broadcast radio has a time constant of 50μs?
10. What is an FM discriminator?
11. Describe the characteristic of an ideal FM discriminator.
12. Name the four FM demodulators other than the PLL.
13. Sketch the block diagram of a phase-locked loop (PLL) FM detector and describe its operation.
14. What is the advantage of the PLL over the over other FM demodulators?

Section 7 Introduction to receivers

This section aims to develop your understanding of the tuned radio frequency (TRF) AM receiver.

At the end of this section you should be able to:

a) Explain the parameters of *sensitivity* and *selectivity* with regard to receivers.

b) State the minimum essential requirements for a basic receiver.

c) State the main disadvantage of the TRF receiver.

Receiver performance parameters

There are several important performance parameters for any receiver. For an introductory subject on electronics communications, we'll consider the most important two: *sensitivity* and *selectivity*.

Sensitivity

The *sensitivity* of a receiver is the input signal voltage required to produce a specific output voltage (say, 1V). For example, a receiver with a sensitivity of 50μV means that it requires 50μV at its RF input (that is, at the input to the electronics from the antenna) to drive the receiver's output to the specified level. It stands to reason that the higher the gain of the receiver, the smaller its sensitivity figure. High performance receivers can have sensitivities down to nanovolts (nV)!

There's more to achieving a desirable receiver sensitivity than simply designing RF amplifiers with high voltage gains. That's because amplifiers introduce noise themselves which gets amplified too. If this noise is larger than the RF signal at the input to the receiver, the signal won't be heard. So, the challenge is to design RF amplifiers so that they have both a high gain and an acceptable signal-to-noise ratio.

Selectivity

The *selectivity* of a receiver describes how capable it is of picking-out the required signal (the "channel" or radio/tv station that you want to hear/watch) and ignoring unwanted signals such as adjacent channels and noise.

If a receiver's selectivity is too low (that is, the tuning circuit's bandwidth is too wide) it will pick up the required signal and may also pick up adjacent channels and/or more noise signals. This will cause the audio output to be a jumbled mix of two or more channels and/or more noisy. See Figure 1a below using broadcast DSBFC AM radio channels as an example.

If a receiver's selectivity is too high (that is, the tuning circuit's bandwidth is too narrow) it may not be wide enough to allow all of the channel's sidebands through, diminishing the quality (fidelity) of the audio signal. See Figure 1b.

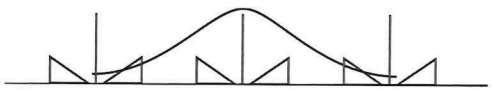


Figure 1a *Under-selective tuning*

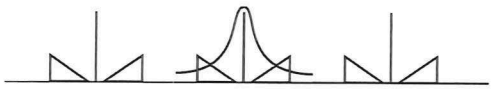


Figure 1b *Over-selective tuning*

The tuned radio frequency receiver

The minimum essential requirements for a radio receiver are surprisingly straightforward and include:

- An antenna – As you know, an antenna is needed to receive the EM signals. (Ideally, the antenna should be efficient for the required band but even inefficient antennas can work sufficiently provided the receiver isn't too far from the transmitter.)
- Tuning – This is needed to "select" the desired channel and reject adjacent channels. Tuning circuits usually also amplify the signal and so active BPFs are usually used.
- Demodulation – This is needed to recover the message.
- Audio amplification – Some power amplification of the recovered message signal is needed so that it can drive a loudspeaker (or headphones).

The tuned radio frequency (TRF) receiver was one of the earliest and simplest receivers to meet these requirements and was used commercially for AM receivers in the 1920s until the development of the superheterodyne receiver. The block diagram for a TRF used to implement an AM receiver is shown in Figure 2.

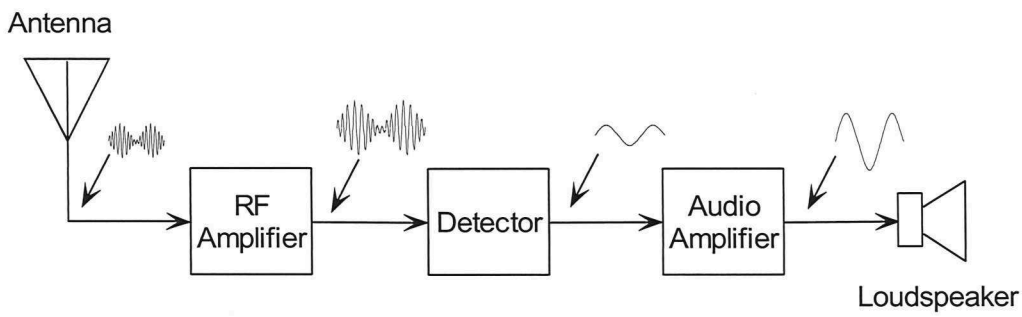


Figure 2 *The block diagram for a simple TRF receiver with expected waveforms*

Tuning is provided by the RF amplifier stage (or stages if there was more than one, which was common). The selectivity of the RF amplifier is essentially set by the selectivity (Q) of its band-pass response. Recall that the Q of an LC band-pass filter is the ratio of its centre (or resonant) frequency to its bandwidth (and has no units). That is:

$$Q = \frac{fr}{BW}$$

If the receiver is required to tune into different channels, the centre frequency of the RF amp's LC circuit can be made adjustable. This can be done by making the capacitor and/or the inductor adjustable.

Limitation of the TRF receiver

The TRF receiver has several limitations. For an introductory subject on communications, we'll consider the main one that contributed to the universal commercial uptake of its successor, the superheterodyne receiver.

For the purpose of the exercise, let's use AM broadcast radio as an example and suppose we design the LC circuit to tune into a channel transmitting on a carrier at 630kHz. As the bandwidth of AM transmissions is 17kHz, the Q of the AM receiver's RF stage must be designed to be 37 (proof: Q = 630kHz ÷ 17kHz = 37.06).

But now let's suppose we tune the same AM receiver to a channel towards the other end of the AM band, say 1557kHz. Either the value of the inductor or capacitor would need to be adjusted but the circuit's Q would also need to change to become 91.59 (proof: Q = 1557kHz ÷ 17kHz = 91.59).

Designing a variable LC tuning circuit that also provides precisely the correct Q at every 9kHz channel allocation between 531kHz and 1602kHz would be a profoundly difficult task to achieve and implement in mass production, if not impossible. Inevitably, designs that permitted variable bandwidths across the band were necessary but this risked the receiver being over or under selective with the problems that go with that.

Another limitation of the TRF receiver – Not tested

An interesting practical limitation of the TRF receiver, especially as a consumer product, turned on the implementation of the RF Stage. To obtain sufficient sensitivity (ie gain at the front end) three cascaded RF amplifiers were often essential and all needed adjusting when the listener wanted to change channels.

In the earliest TRF receivers, the RF amplifiers had to be adjusted manually, one at a time (which involved some toing-and-froing between the three stages) that made tuning challenging. Later models employed ganged tuning components for the RF amplifiers but "tracking errors" between them still required fine adjustments to correct.

The superheterodyne receiver solved both this problem and the also the variable selectivity described above.

Review questions

Answer these questions to check your understanding of what you have learnt from this chapter. Doing this will help you prepare for the knowledge assessment.

1. Describe the term *sensitivity* with respect to a radio receiver.
2. Describe the term *selectivity* with respect to a radio receiver.
3. What would be the effect of a low selectivity on a receiver?
4. What would be the effect of an overly high selectivity on a receiver?
5. State the five requirements of a basic receiver.
6. What is the main limitation of the TRF receiver?

Student notes

Section 8 The superheterodyne receiver

This section aims to develop your understanding of the operation of the superheterodyne receiver and its advantage over the TRF receiver.

At the end of this section you should be able to:

a) Sketch the block diagram of a single conversion superheterodyne receiver that can be used to receive of AM/FM broadcast radio signals.

b) Explain the purpose of each block in a single conversion superheterodyne receiver.

c) Define the term *intermediate frequency*.

d) Explain how the superheterodyne receiver selects the desired AM channel (on a given frequency) and list the relevant frequencies at the output of each block.

e) Explain how the superheterodyne receiver rejects adjacent channels.

f) State the intermediate frequencies used for Australian AM and FM broadcast radio.

g) Define the term *image frequency* with respect to the superheterodyne receiver.

h) Calculate a superheterodyne's image frequency for a given desired channel frequency.

i) State the main advantages of the superheterodyne receiver over the TRF receiver.

j) State the purpose of the automatic gain control (AGC) and show how this stage fits into the block diagram for a superheterodyne receiver.

k) Verify the operation of the superheterodyne method of selecting (tuning into) desired channel frequencies.

The superheterodyne receiver

The superheterodyne (or "superhet") receiver was invented in the 1930s to overcome the limitations of the TRF receiver discussed in the previous chapter and has been widely used ever since in all sorts of applications but especially radio communications (both, AM and FM).

Figure 1 below shows the basic block diagram of a superheterodyne receiver used to receiver AM broadcast radio signals.

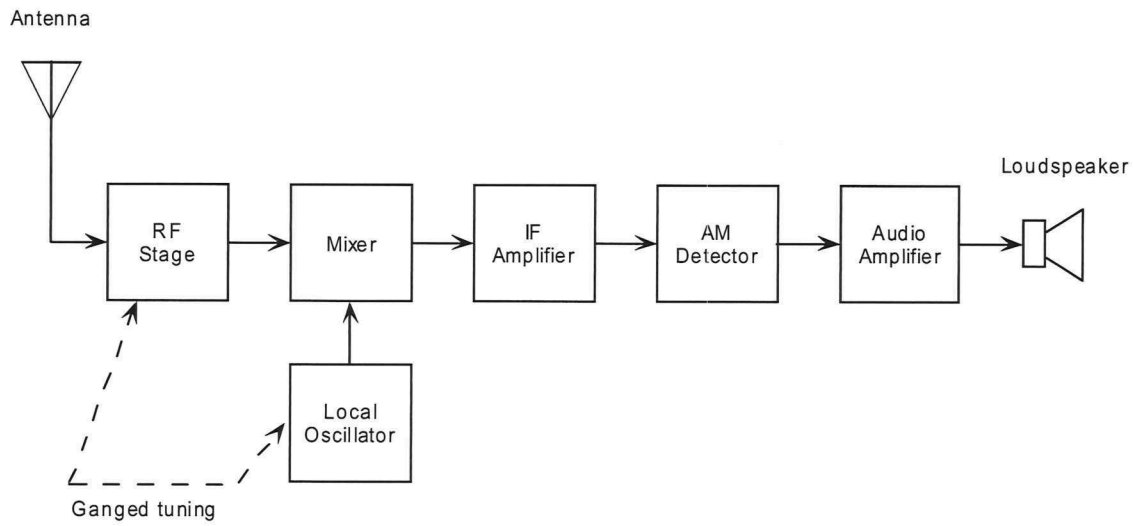


Figure 1 *The block diagram of a basic AM superheterodyne receiver*

The diagram in Figure 1 is "basic" because it shows the minimum blocks the superhet receiver needs to be implemented. (Actually, even the RF stage isn't necessarily needed so this could have been left off too.) It's often the case that receiver designs include additional functional blocks to improve the performance of the receiver depending on the communications system's requirements but these extra blocks are not relevant to this discussion.

Operation
The following notes explain the operation of the superhet receiver using the example of an AM broadcast radio receiver for simplicity of explanation.

You already know the purpose of the **antenna**. It's a transducer that converts EM transmissions into electrical signals. The magnitude of the signal out of the antenna is likely to be in the order of microvolts.

The **RF Stage** is typically a tuneable active band-pass filter (though some receiver designs implement the filtering and amplification as discrete cascaded circuits). Importantly, although the RF Stage is tuned to the carrier frequency of the selected channel, its bandwidth is fairly wide and so the signal from one or more adjacent channels will also appear on its output. In other words, the RF Stage does provide the receiver's selectivity (as it does in the TRF receiver).

For superhet receivers that include an RF Stage, the stage provides one or more of the following: a) band-limiting to reduce the noise that enters the receiver in the first instance; b) an improvement in the receiver's sensitivity; and c) "image rejection" (more later).

The **Local Oscillator** (LO) is a sinusoidal oscillator with an output frequency that is variable. Importantly, the LO's output frequency control is "ganged" with the RF Stage's tuning – This means that they're adjusted together. Equally importantly, for Australian AM broadcast radio, the LO's output frequency is always 455kHz higher than the frequency of the channel that the RF Stage is tuned to.

For example, suppose the receiver is being used to listen to a radio station (that is, a channel) on 630kHz. The centre frequency of the RF Stage is 630kHz and the output frequency of the LO is 1085kHz (ie 630kHz + 455kHz).

The **Mixer** is a non-linear circuit that accepts two inputs; the output of the RF Stage and the sinusoid from the LO. The Mixer non-linearly combines the two signals to produce a set of sum and difference frequency components (plus a bunch of other mostly insignificant signals too but these can be ignored for the moment).

Using the example of the receiver tuned to an unmodulated sinusoidal carrier transmitted on 630kHz (and so the LO's output frequency is 1085kHz), the Mixer's output includes two sinewaves at the sum and difference frequencies of:

- 1715kHz (ie 1085kHz + 630kHz)
- **455kHz** (ie 1085kHz – 630kHz)

If the 630kHz carrier is modulated, the carrier will have sidebands (ie sinewaves straddling the carrier). This means that the two new sinewaves on the Mixer's output (at 455kHz and 1715kHz) will also have sidebands. For example, suppose the 630kHz AM carrier is modulated by an 8.5kHz sinusoidal message. Using math you've seen before, the AM signal consists of three sinewaves, 621.5kHz (f_{LSB}), 630kHz (f_C) and 638.5kHz (f_{USB}).

So, when this signal is mixed with the LO's 1085kHz sinewave, the resulting spectral composition of the Mixer's output includes the sum and difference frequencies of the LO's output for all components in the AM signal. For our 630kHz carrier amplitude modulated by the 8.5kHz sinusoidal message, the Mixer's output includes the following sets of sinewaves:

- 1706.5kHz, 1715kHz and 1723.5kHz
- 446.5kHz, **455kHz** and 463.5kHz

These spectral components on the Mixer's output are shown in Figure 2 below.

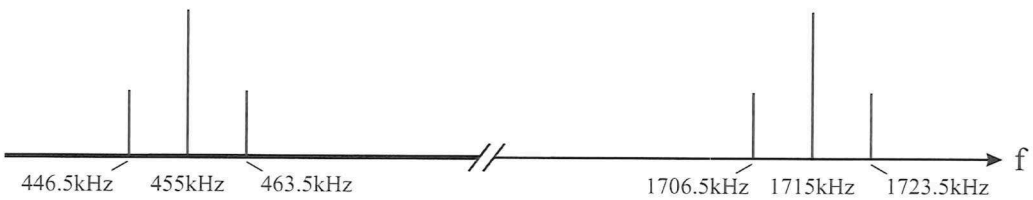


Figure 2

It can be seen from the math and/or the drawing in Figure 2 that, by mixing an AM signal with the LO signal, two completely new AM signals are generated! Importantly, these two AM signals are carrying the exact same information that was in the original AM signal (that the receiver is tuned to), except on different carrier frequencies; one at a much higher frequency than the original carrier and one at a lower frequency.

Frequency translation – Not tested
Textbooks and websites sometimes refer to this process as "frequency translation". Two other terms that are used are "down-conversion" and "up-conversion" with the choice being determined by which of the two new AM signals is wanted. As you will see, in an AM superhet radio receiver, it's the one on the 455kHz carrier that is wanted so this would process would be referred to as "down-conversion".

Now, the two AM signals on the Mixer's output are then fed to the **IF Stage** which is an active band-pass filter with a centre frequency of 455kHz and a bandwidth of 17kHz.

Given the IF Stage's performance specification quoted in the previous paragraph, it's clear that the AM signal with the 455kHz carrier frequency on the Mixer's output would pass through the IF Stage and on to the next stage. All other signals on the Mixer's output would be blocked (rejected). This is shown in Figure 3 below.

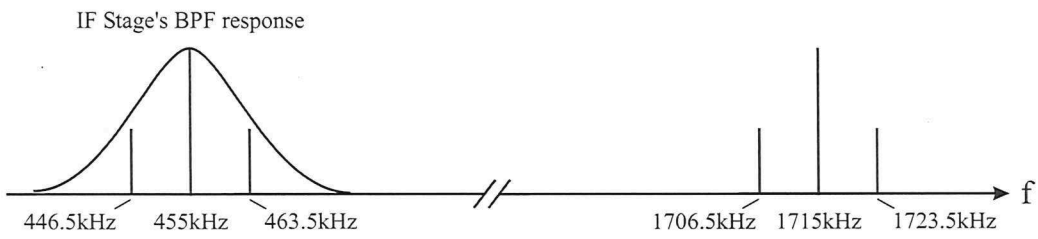


Figure 3

As you already know, the **AM detector** demodulates AM signals and so it would demodulate the IF Stage's output signal to recover the original message signal. And the **audio amplifier** provides the required power gain to drive the loudspeaker which converts the electrical signal to sound pressure waves so they can be heard.

Tuning

So, how does the superhet tune into different radio stations (channels)? As already mentioned, when tuning into a channel, the RF Stage's centre frequency is adjusted to be the same as the channel's carrier frequency but the LO's output is set so that its frequency is 455kHz.

The previous discussion explained the operation of the superhet using the example of tuning into a channel on a carrier of 630kHz. To see how other channels can be selected, let's redo the calculations for the same receiver tuned into a channel on a carrier of 1557kHz that is also amplitude modulated by a 8.5kHz sinusoidal message.

To tune into that channel, the RF Stage's centre frequency must be 1557kHz (ie the same frequency as the desired channel's carrier frequency) and the output frequency of the LO must be 2012kHz (ie 1557kHz + 455kHz). The resulting significant sinewaves on the Mixer's output due to the desired channel includes two new AM signals:

- 3560.5kHz (f_{LSB}), 3569Hz (f_C) and 3577.5kHz (f_{USB})
- 446.5kHz (f_{LSB}), **455kHz** (f_C) and 463.5kHz (f_{USB})

Did you see what happened there?

Tuning into a channel on a carrier frequency of 1557kHz produces an AM signal with a carrier of 455kHz, including sidebands containing the information in the original message, just as the superhet did when it was tuned to a channel on a carrier of 630kHz.

In other words, no matter what channel the receiver is tuned to within the band, a copy of that channel's AM signal is always generated on a 455kHz carrier which is then passed to the detector to be demodulated by the IF Stage.

You can possibly see how the IF Stage gets its name now. The word "intermediate" means, "coming between two things". This is exactly what is happening with the IF signal in the superhet receiver - There is no such signal in the TRF receiver because it detects (ie demodulates) the received AM signal of the desired channel, not a "down-converted" version of it.

Selectivity

By now you're probably wondering what the point is of this elaborate intermediate step in the process of demodulating the channel.

Recall that one of the TRF's problems (it has several) is the fact that all of the tuning is implemented by the BPF in the RF Stage and it is difficult (if not impossible) to design it so that it has exactly the right Q-factor across the band to achieve a fixed 17kHz bandwidth. This adversely affects the receiver's selectivity rendering it either under or over-selective for some channels.

How does the superhet solve this problem?

You have seen that, regardless of the desired channel's carrier frequency, the superhet always down-converts the desired channel's modulated signal (AM or FM) to a modulated signal on 455kHz. Importantly, this means that the IF Stage doesn't need to be designed to be adjustable to tune into different carrier frequencies. So, its centre frequency is fixed which, in turn, means that its bandwidth is fixed too, at 17kHz (giving an $f1$ at 446.5kHz and an $f2$ at 463.5kHz). Problem solved!

To see this selectivity in action in terms of tuning out (that is, rejecting) other channels, let's return to the receiver tuned to an AM channel on 630kHz which, as you know, the superhet down-converts to an AM signal on 455kHz. Let's also assume there is an adjacent AM channel transmitting on 648kHz (the closest an adjacent channel can be) having been modulated by an 8.5kHz sinusoidal message.

As the RF Stage has a wide bandwidth, this adjacent AM channel's signal passes through the RF Stage to the Mixer's input as well as the channel on 630kHz. This means that the adjacent channel's signal is also mixed with the LO's signal.

Recall though, as the RF Stage is tuned to the AM channel on 630kHz, the LO's output frequency is 1085kHz. So, the adjacent channel on 648kHz is mixed with 1085kHz producing two AM signals having the following spectral components:

- 1724.5kHz (f_{LSB}), 1733Hz (f_C) and 1741.5kHz (f_{USB})
- 428.5kHz (f_{LSB}), **437kHz** (f_C) and 445.5kHz (f_{USB})

The down-converted version of the AM signal on 648kHz is shown in Figure 4 together with the down-converted version of the desired channel on 630kHz.

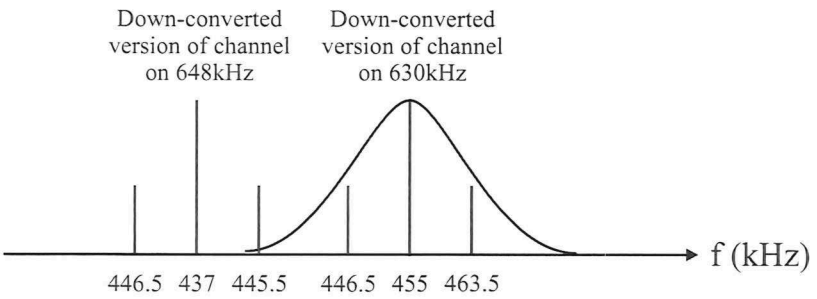


Figure 4

It's clear from this drawing - which also shows the IF Stage's response on the spectrum - that the down-converted version of the adjacent channel on 648kHz can't pass through the IF Stage. Ideally, only the desired channel's signal is down-converted to frequencies that can pass though the IF Stage.

To improve the selectivity of the IF Stage (and hence further improve the selectivity of the receiver) a combination of one or more of the following design techniques are employed: 1) The use of cascaded filters (and so the IF Stage is often referred to as the "IF Strip") which is easily implemented because of its fixed-frequency tuning; and/or 2) The use of ceramic or crystal filters. These give the stage's BPF response a much sharper roll-off at each end rejecting as much as possible of the adjacent down-converted signal's sidebands to prevent them from slipping through the IF Stage and causing interference.

The IF Stage also provides the majority of the superhet receiver's sensitivity too (or the entire sensitivity where an RF stage isn't used). This is because the signal on the Mixer's output is still relatively small and needs further amplification prior to demodulation which the IF Stage implements.

The superhet and the problem of "image signals"

There is a problem with the superheterodyne receiver. It is possible for an unwanted channel to be converted to a modulated signal 455kHz and pass through the IF Stage together with the desired channel's signal. This unwanted channel's carrier is known as the "image frequency".

To explain how image signals are generated, let's return one last time to our AM superhet receiver tuned to a channel transmitting on 630kHz. Let's also assume there is channel transmitting on 1540kHz having been modulated by an 8.5kHz sinusoidal message.

Again, as the RF Stage is tuned to the AM channel on 630kHz, the LO's output frequency is 1085kHz. So, when mixed at the receiver, the channel on 1540kHz produces two AM signals having the following spectral components:

- 2616.5kHz (f_{LSB}), 2625kHz (f_C) and 2633.5kHz (f_{USB})
- 446.5kHz (f_{LSB}), **455kHz** (f_C) and 463.5kHz (f_{USB})

If you double-check the math for yourself, you'll see that that the AM signal on 1540kHz produces an AM signal on 455kHz when 1085kHz is subtracted from its spectral components. So, this signal will pass through the IF Stage as well as the desired channel's down-converted signal! The two message signals will be heard at the receiver's output and will obviously interfere with each other.

The image frequency is always equal to the desired channel's carrier frequency plus twice the IF Stage's frequency which can be expressed as the equation:

$$\boxed{f_{image} = f_C + 2f_{IF}}$$

Where f_C = the channel's carrier frequency and f_{IF} = the IF Stage's frequency.

One method of managing the problem of image signals is to assign channel frequencies to prevent users in the same reception area from transmitting on another channel's image frequency. For example, if there is a channel transmitting on 630kHz then its image frequency (1540kHz) would not be allocated to another user. Conversely, if there is a channel on 1540kHz then its image frequency (630kHz) would not be allocated to another user.

However, managing the problem this way makes for inefficient use of the band's spectrum. Moreover, image signals may exist in a lower or higher band allocated for other purposes which makes for potentially inefficient use of those bands instead, or as well. And then there's the problem of potential image signals from afar by travelling long distances at night via inadvertent sky wave propagation.

An alternative solution to managing image problems is for the receiver to include an RF Stage that is designed with a bandwidth that rejects images from getting through to the Mixer. Although the RF Stage's bandwidth is typically wide, it usually designed to be narrow enough to prevent image signals from reaching the Mixer.

Having said that, this solution may not work for communications systems based on higher frequency carrier signals and using a relatively low IF frequency (like broadcast FM). The RF Stage's bandwidth may not be able to be made sufficiently narrow to reject images without introducing the problems associated with the TRF receiver as the receiver is tuned across the band.

Double conversion – Not tested

Possibly the best solution for preventing image problems is a technique known as "double conversion", though this can increase the receiver's manufacturing cost. In double conversion systems, the desired channel's carrier is down-converted to a new modulated carrier that passes through the first of two IF Stages.

This signal is then down-converted a second time (using a second Mixer and LO combination) to a second lower carrier frequency that passes through the second IF Stage tuned to that lower frequency. As this is an advanced communications design, double conversion is not tested and so it's not explained any further here.

If you're interested to know more about double conversion receivers, visit:

https://www.electronics-notes.com/articles/radio/superheterodyne-receiver/double-triple-dual-multiple-conversion.php

Automatic gain control (AGC) in the supehet

All AM receivers incorporate an *automatic gain control* (AGC) circuit that adjusts the gain of the receiver depending on the strength of the received signal. If the received signal is strong, the AGC reduces the receiver's gain. If the received signal is weak, the AGC increases the receiver's gain. The AGC, therefore, provides constant speaker "loudness" regardless of the received signal strength, the location of the receiver and the weather and ionospheric conditions.

To explain, without AGC, the user would have to turn the volume up when tuning into weak signals and down for strong signal. And they would have to turn the volume up as they travel away from the transmitter and down when travelling towards it. And the volume would likely also needing adjusting throughout the day and week as a result of changes in the weather and ionospheric conditions.

In most AM superhet receivers, the signal to help set AGC is obtained at the output of the detector. An RC filter with a relatively long time constant produces a DC control voltage proportional to the received signal strength. This control voltage is fed back to the IF and/or RF stages to control their gain.

The block diagram for an AM superheterodyne receiver with AGC is shown in Figure 5.

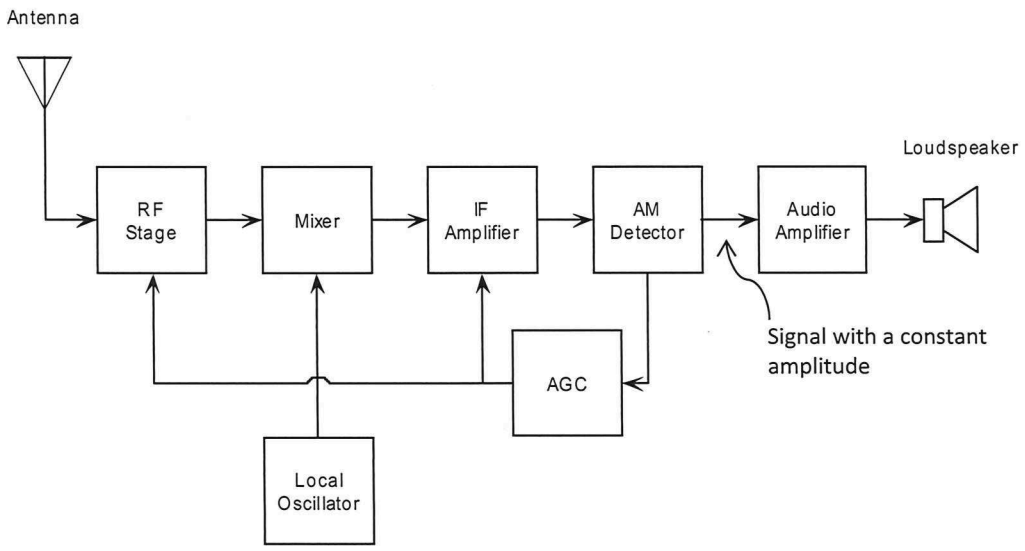


Figure 5 *The block diagram for an AM superheterodyne receiver with AGC*

The superheterodyne receiver for broadcast FM radio

The FM broadcast radio superheterodyne receiver works on exactly the same principle as AM superheterodyne receiver discussed previously.

There are several implementation differences though. First, the FM superhet tunes across the FM broadcast band (approximately 88MHz to 108MHz) instead of the AM broadcast band. Second, the FM superhet uses an IF of 10.7MHz instead of the 455kHz figure used for AM. Third, the FM superhet typically includes a limiter stage (to provide a form of AGC even if the limiter isn't needed to improve the receiver's noise performance). Fourth, the FM superhet uses an FM discriminator or PLL instead of an AM detector. Fifth, the FM superhet includes a de-emphasis stage to correct for the pre-emphasis applied to the original message signal at the transmitter.

The block diagram for an FM superheterodyne receiver is shown in Figure 6 below.

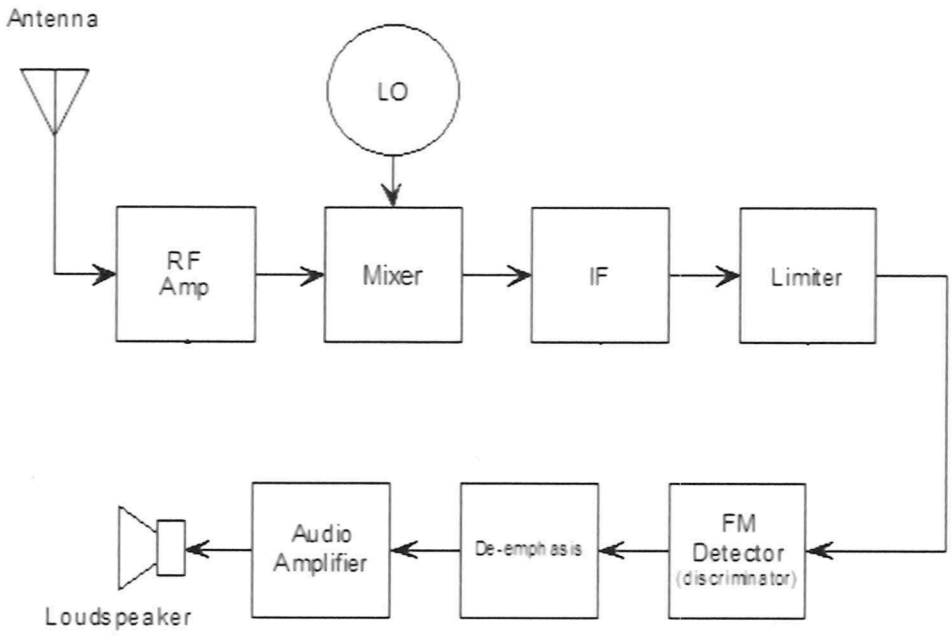


Figure 6 *The block diagram for an FM superheterodyne receiver*

AGC in FM receivers – Not tested
Although AGC is not essential for FM superheterodyne receivers with a limiter, AGC can be included to avoid problems caused by radio signals that are too large and overload the receiver's front end (ie the RF stage). This is done using a circuit called the RSSI detector to create the control voltage needed to affect the gain of the RF stage.

The future

With digital message signals being used more and more, it's possible that the superhet's days are numbered as the near universal method for implementing receivers.

"Direct conversion" of communications signals down to baseband frequencies using a process called "undersampling" is increasingly being used together with demodulation that is implemented in software (rather than circuits like the envelope detector) to create a receiver technology called "software defined radio" (SDR). If you work with communications, you may well have heard of this technology. If you haven't, then you will.

Review questions

Answer these questions to check your understanding of what you have learnt from this chapter. Doing this will help you prepare for the knowledge assessment.

1. Explain the main difference between a TRF receiver and a superheterodyne receiver.
2. Sketch the block diagram for an AM superheterodyne receiver showing the following blocks and direction of signal flow:

   - Audio amplifier
   - AGC
   - Loudspeaker
   - Mixer
   - Detector
   - Antenna
   - RF Amplifier
   - Local oscillator
   - IF amplifier

3. Describe the function of the following blocks:
   - Mixer
   - Local oscillator
   - IF amplifier
   - AGC
4. What is the centre frequency of the IF amplifier in broadcast AM radio receiver?
5. What is the centre frequency of the IF amplifier in a broadcast FM radio receiver?
6. Explain why a superheterodyne receiver has superior selectivity compared to a TRF receiver.
7. Calculate the local oscillator frequency in an AM superheterodyne receiver tuned to a channel on 1134kHz.
8. If the carrier in the previous question was modulated by a 5kHz sinewave, calculate the frequency of the six significant sinewaves at the output of the mixer due to this channel.
9. Describe what is meant by the term *image frequency*?
10. Explain how an *image frequency* can get through the IF stage.
11. Calculate the image frequency being received on a conventional broadcast receiver that has an IF of 455kHz when the receiver is tuned to 1134kHz.
12. How does the RF Stage reduce interference caused by image frequencies?
13. Explain the general concept of AGC in an AM receiver and why it is needed.
14. Explain how the AGC control voltage is generated and to which stages it is fed.

15. Sketch the block diagram for an FM superheterodyne receiver that includes the additional following blocks and direction of signal flow:

    - Pre-emphasis
    - Limiter
    - FM discriminator

16. Explain the purpose of the *limiter* in an FM superheterodyne receiver.

17. Explain why AGC is not necessarily essential in FM superheterodyne receivers than include a limiter.

18. Calculate the typical frequency range of the local oscillator in a broadcast FM receiver (assuming the FM band to be 88.1MHz to 107.9MHz)?

19. Calculate the local oscillator frequency in an FM superheterodyne receiver tuned to a channel on 91.3MHz.

20. Calculate the image frequency that could cause interference to the channel on 91.3MHz.

Student notes

Section 9 Optical communications principles

This section aims to develop your understanding of basic optical communications systems principles, media, and advantages.

At the end of this section you should be able to:

a) State the two materials used to implement optical fibres.

b) List the advantages of transmitting information over optical fibres compared to using metallic conductors and EM waves in free space (ie wireless).

c) Appreciate that fibre-optic communications systems can be represented by the generic block diagram of a communications system.

d) Name and briefly describe the optical principle that permits the propagation of light along an optical fibre.

e) Name the two types of cable connections/terminations used for fibre-optic cables.

f) List the hazards when working with fibre-optic communications systems and state the appropriate control measures.

Fibre-optic communications systems

Introduction

Optical fibres conduct light in place of electrical signals and were first developed for commercially use in the 1970s. Since then, they have steadily replaced many copper wire communications systems.

Today, fibre-optic cable communications systems are used extensively in telecommunications to provide telephone and broadband internet. In fact, fibre-optic systems helped to kick-start the "information age" that we now enjoy. (The internet has been around since the 1960s but your author only used it for the first time in late 1992 and only for email. Apparently, at the end of 1992, there were only 26 websites. The world wide web didn't really take off until about 1994.)

An example of glass fibre-optic cable used for telecommunications and consisting of hundreds of fibres is shown in Figure 1.

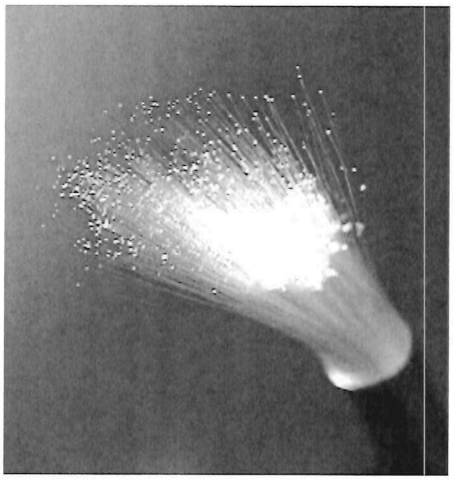


Figure 1 *Fibre-optic cable (Ref: http://www.cablexpress.com/blog/?p=183)*

(At this point, a distinction should be made between two key terms used throughout this chapter. "Optical fibre" is the individual fibre along which the light travels. "Fibre-optic cable" is the entire cable consisting of one or more optical fibres, surrounded by buffers, strength membranes, armour, jacket, etc.)

Fibre-optic cable communications are not limited to telecommunications and broadband internet. They're used today in many other applications such as in military equipment, medical equipment, control equipment, and even in every-day applications like home networks and car networks (with plastic more likely being used for this last example).

The main advantages of fibre optic cable communications systems

Optical fibre has numerous advantages over copper and other metallic conductors used for transmitting signals. The main advantages include:

- Much lower signal attenuation

  EM signals transmitted over metallic conductors lose much more of their signal strength over any given distance compared to light transmitted over glass optical fibres. This means that fewer "repeaters" (amplifiers) are needed for long distance (or "long haul") fibre-optic cable communications than for the same distances spanned by copper wire communications.

  It's this very issue that gave fibre-optic cables and systems their start because they were incredibly expensive at first (like most new technology). Underwater cables (called submarine cables) providing telephone communications between countries and continents needed repeaters at set distances to avoid losing the signal due to attenuation. Switching to fibre-optic cables and systems increased the distance between repeaters thus reducing their number with obvious engineering and cost benefits considering they are located at the bottom of the sea and ocean and so difficult to reach. Developments in fibre-optic submarine cables together with economies of scale made fibre-optic cables generally more advantageous to use and they quickly also started replacing land-based long-haul (or "backbone") telecommunications cables (ie cables between major cities, like Sydney and Melbourne).

- Extremely wide bandwidths permitting extremely high data rates

  Depending on the design of the metallic transmission line, data rates range from about 100Mbps to 1Gbps. Optical fibre data rates can exceed 10Gbps. Moreover, by using wavelength-division multiplexing (WDM) data rates can exceed 16Tbps.

  As an aside, the issue of bandwidth was not as compelling an advantage for using long-haul fibre-optic cables as it is today. The telephone system up to the 1980s was designed for the constrained bandwidths of twisted-pair cables and so the traffic was relatively light by today's standards. Long distance voice communications was expensive (your author remembers calling family in the UK around in 1990 and it costing around $2/minute which is about $10/minute in today's money) and so people were more likely to send letters unless the communication was urgent. Moreover, the overwhelming majority of data on the lines came from telex machines, then facsimile (fax) machines (which started replacing telex from the late 1980s) which have relatively low data requirements.

  As fibre-optic cables replaced copper long-haul cables, it created opportunities to send more data for the same cost which fuelled a demand in exchanging more data. This process started a virtuous circle so that now we have fibre-to-the-node (FTTN) telecommunications here in Australia implemented by NBN Co. For most people in Australian cities, only the so-called "last mile" from the node to your home is copper.

- Immunity to electromagnet interference (EMI from EM radiation) at radio frequencies
  Although light is an electromagnetic wave, it is largely unaffected by EM waves at radio frequencies. In fact, light is largely unaffected by other light. For a practical example of this, you can cross the beam of two torches without noticing any effect on either of the beams on the surface that they strike. (That said, this doesn't mean that light beams don't interact in the region where they cross but that's a different thing.) This means that fibre-optic cable communications are not affected by naturally occurring or human-made sources of EMR like wireless and other cable communications.

- No crosstalk
  "Crosstalk" is a problem with adjacent metallic conductors. The electromagnetic radiation from each cable is picked-up by the other, converted to electricity and added to the electrical signal. This can cause interference. Light can't escape optical fibres and fibre-optic cables are shrouded and so no light can enter them either. Moreover, light is largely unaffected by other EM waves. So, there is virtually no possibility of crosstalk between adjacent optical fibres or adjacent fibre-optic and metallic cables.
  (That said, crosstalk can be a problem when optical fibre is being used for simultaneous bi-directional communications using wavelength division multiplexing due largely to impurities in the fibre and the practical limitations of coupling light to the fibre at each end.)

- More secure communications
  Metallic conductors are easy to "tap" and a direct electrical connection is not even needed (a simple loop of insulated wire around the conductor can serve as an effective tap). Tapping optical fibres is much more difficult to do and easily noticed at the receiver due to the significant drop in light signal strength that it would cause.

Fibre-optic communications also has advantages over wireless for point-to-point communications (ie from one transmitter to one receiver) including:

- Higher data rates (for now, at least)
- Immunity to EMI caused by EMR
- Immunity to changes in the weather and ionosphere
- More secure

The fibre-optic cable communications system

Fibre-optic cable communication systems are really just like any other communications system in most respects. A message signal is used at the transmitter to modulate a carrier which is transmitted over a medium, then converted back to a message signal at the receiver.

However, the difference between fibre optic communication systems and copper/wireless communications is two-fold. First, the carrier is light instead of EM waves in the radio spectrum. Second, the medium is a glass or clear plastic conductor instead of copper conductors or free space.

Figure 2 below shows the main components in a fibre-optic communications system. You'll notice that it's effectively most of the generic block diagram of a communications system taught to you at the beginning of this Unit (see Section 1). The input signal comes from the source. The "modulator" and "light source" (a laser or an LED) together implement the transmitter. The optical fibre is the medium. The "light detector" and "demodulator" together implement the receiver. The output signal goes to the destination.

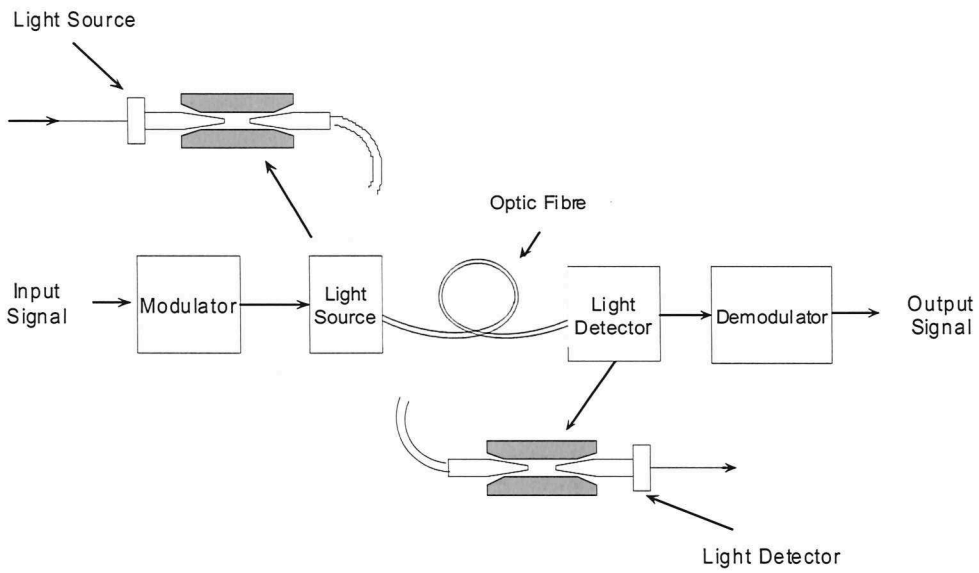


Figure 2 *The main components in a fibre-optic communications system*

How it works – Not tested

The input signal in Figure 2 is the message signal. The message in a fibre-optic communications system can be analog or digital. However, it's much more likely to be digital. That's either because this is how the information exists in the first place (eg text in an email). Or because message signals like speech or music are converted to digital data first which permits the use of digital techniques like data compression, error checking, etc. A variety of digital modulation schemes can be used with fibre-optic communications and pulse-code modulation (PCM) is widely used in telecommunications.

Once the message has been converted to digital, the signal is used to switch the light source on and off (or between two light levels). The light source could be a device as simple as a visible/infrared light LED (used in short-run plastic fibre systems) or a laser source (used in telecommunications systems).

The light source is then coupled to the optical fibre inside the fibre optic-cable. This coupling can be as simple as a mechanical connector as shown in Figure 2 or by more sophisticated means like the devices used in telecommunications called *circulators* which introduce fewer losses.

The light is then propagated along the optical fibre due to *total internal reflection* (more later) and reaches the other end as an attenuated light signal and possibly containing errors due to the limitations of implementing the fibre itself (impurities in the glass) or problems caused by the installation (the need to use connectors/splices, using connectors with losses that are too high, poorly executed splices, bends in the cable that are too tight, etc).

The light is decoupled from the fibre using a mechanical connector or circulator and a photo-sensitive device is used to return it to an electrical signal. After that, the signal is demodulated to return it to its original electrical form ready for use at the destination.

Principles of fibre-optic light transmission

The propagation of light along optical fibre relies on a phenomenon called *total internal reflection* which needs a little explaining.

Figure 3 below shows two materials each having different refractive indexes (n) with Material 1 having a higher refractive index than Material 2. The diagram shows light from inside Material 1 striking the interface between the materials at an angle that is perpendicular to the interface. Assuming the interface is perfect (ie the two surfaces are perfectly smooth and clean), all of the light passes (is transmitted) from Material 1 to Material 2.

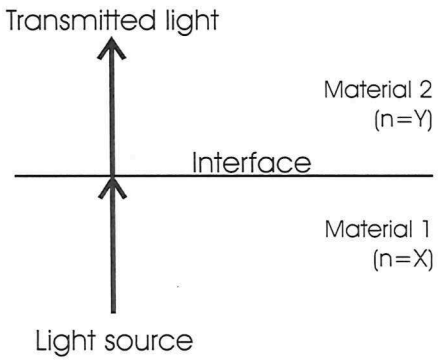


Figure 3

Now, suppose the angle of the light from inside Material 1 strikes the interface at a small angle to its original position as shown in Figure 4. Most of the light is transmitted into Material 2 with some refraction taking place but some of the light is also reflected back into Material 1. (Note: Line thickness has been used to convey the approximate relative distribution of light energy that is transmitted and reflected.)

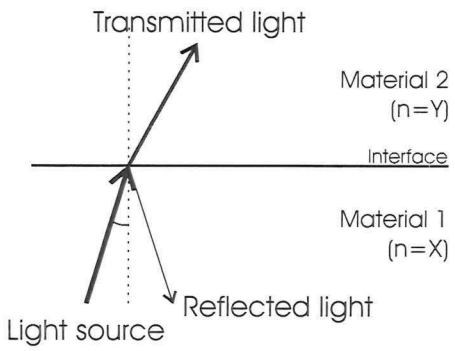


Figure 4

Importantly, as the angle of the light from inside Material 1 increases from the perpendicular, less light is transmitted into Material 2 and more is reflected back into Material 1. This is shown in Figure 5.

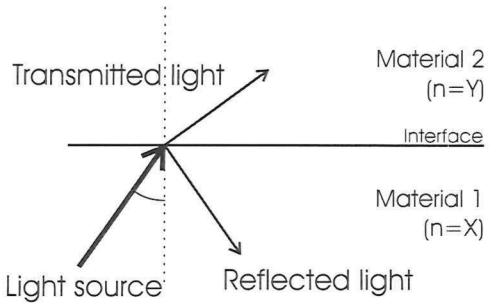


Figure 5

When the angle of the light from inside Material 1 reaches and exceeds an angle called the "critical angle" no light is transmitted into Material 2 and all of the light is reflected back into Material 1. This is *total internal reflection* and is shown in Figure 6.

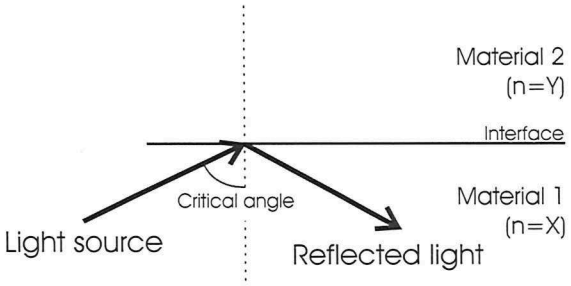


Figure 6

The refracted and critical angles can be calculated if the materials' refractive indexes are known but that's not necessary for this Unit. An important attribute that you must know for this Unit is that a light wave's angle of incidence and angle reflection are always the same (just like a ball in snooker when it strikes one of the cushions!).

Now consider the construction of a simple fibre-optic cable in Figure 7 below. The optical component of the cable consists of the solid ultra-pure glass core (or plastic for cheaper short-run cables) with a certain refractive index that is surrounded by "cladding" which is also glass (or plastic) having a lower refractive index.

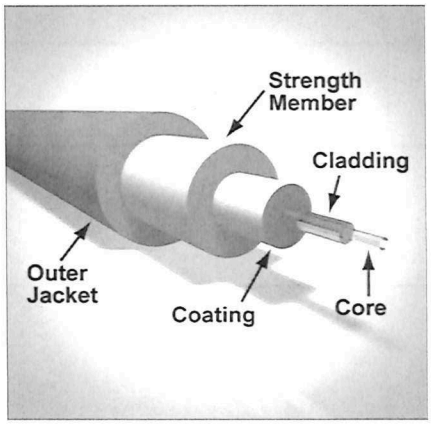


Figure 7 *The construction of a simple fibre-optic cable*

Viewed in cross-section along the length of the optical fibre, the core and cladding look as shown in Figure 8 below.

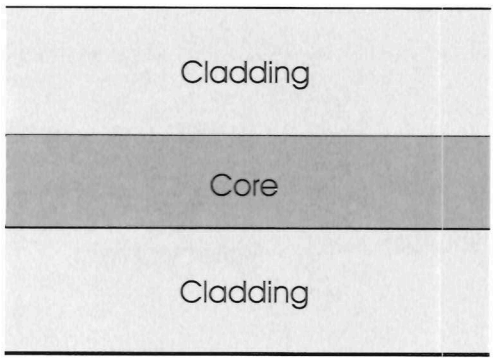


Figure 8

When light is coupled to the input of the optical fibre so that it enters at, or above, the critical angle (for the core-cladding interface), it is reflected down the core due to total internal reflection together with the fact that the light wave's angle of incidence and angle of reflection are always the same. This is shown in Figure 9 below.

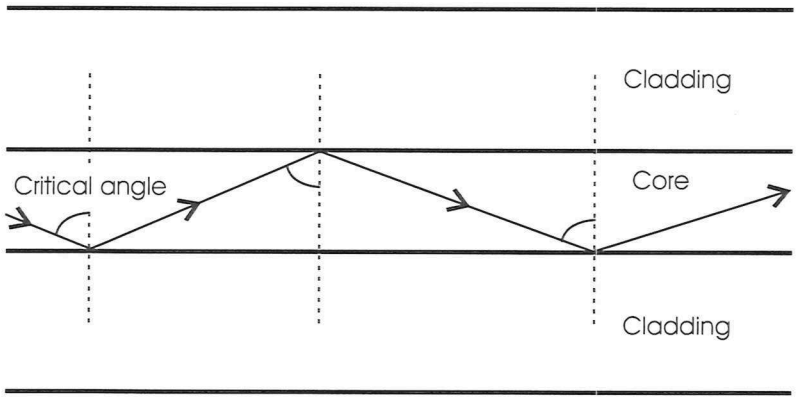


Figure 9

The example of optical fibre shown in Figure 8 is known as a stepped-index fibre. A variation on the idea is the graded-index fibre where there is only a core (no cladding) that has a refractive index that reduces evenly from the centre to its outer edge in all directions achieving the same effect.

Another difference between optical fibre types – Not tested

Another important difference between optical fibre types turns on whether it is "single mode" or "multi-mode". Multi-mode fibre (MMF) permits multiple light rays to enter the fibre at different angles and so light at each angle travels along different paths or modes. Single mode fibre (SMF) permits light rays to enter the fibre at only one angle and so they travel in only one path or mode.

Multimode optical fibres have core diameters from about 50μm to 100μm. Single mode fibre cores have much smaller diameters, typically around 8μm to 10μm. The diameter of the core and cladding together is typically 125μm for both SMF and MMF cables. (To put these dimensions into perspective, human hair varies widely in thickness but a figure roughly in the middle of the range is about 100μm.)

These two attributes of optical fibres (their index and mode) mean that four fibre-optic cable combinations are available:

- Single-mode stepped-index
- Single-mode graded-index (not widely used)
- Multi-mode stepped-index
- Multi-mode graded-index

Each of the four types has advantages and disadvantages compared to the others in terms of cost and performance. If you're interested to know more, visit: https://community.fs.com/blog/single-mode-cabling-cost-vs-multimode-cabling-cost.html

The outer jacket of fibre-optic cables are colour-coded to indicate whether they are single-mode (yellow jacket) or multi-mode (orange jacket or aqua jacket for laser optimised cables).

Termination and splicing

Optical fibres are connected to equipment, or to each other, using either special fibre-optic cable connectors or a joining process called *splicing*.

Fibre-optic connectors

Some fibre-optic connectors are shown in Figure 10. The variety of connectors reflects the variety of cable types mentioned earlier, cost, ease of termination and performance (like the losses introduced by the connector due to the imperfect match between the faces of the optical fibres they join).

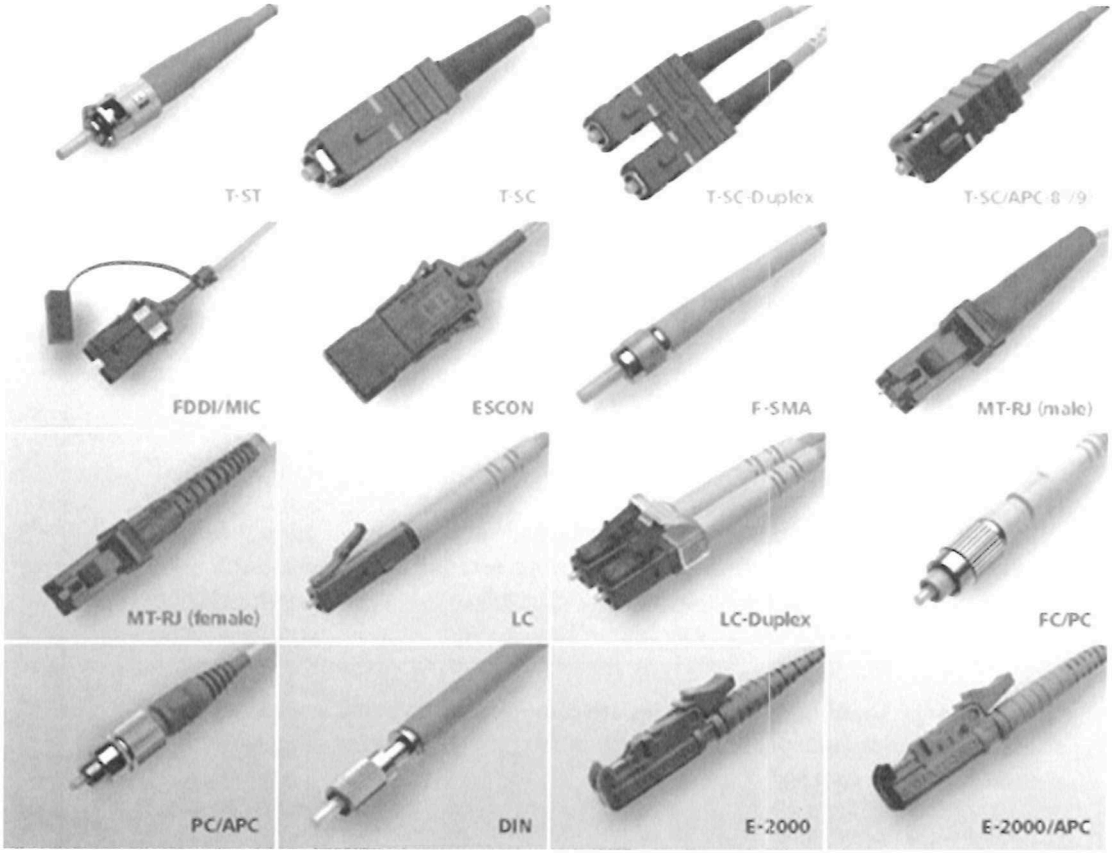


Figure 10 *Fibre-optic connectors*

Splicing

Splicing is a word to describe the joining of two lengths of rope together - without using a knot - so that they effectively become one length of rope. It's a perfect metaphor for joining optical fibre (without using connectors) and there are two types: fusion splicing and mechanical splicing.

Fusion splicing uses heat (typically from an electric arc) to join the two fibre ends together so that the two fibres become one.

Mechanical splicing doesn't join the two fibres like fusion splicing. Instead, the two ends of the fibre are held together face-to-face using a mechanical sleeve.

Both methods require considerable preparation of the ends of the fibre and considerable amount of skill to execute.

Splicing is less bulky than using connectors. And, done properly, splicing results in lower losses across the join and are unlikely to deteriorate over time. Together these benefits make splicing the preferred option for long term joints between optical fibres.

Hazards associated optical fibres

The main hazards associated with fibre-optic cable communications relate to the light source, glass shards from broken pieces of optical fibre, and the chemicals used in splicing.

Light sources used in most fibre-optic cable communications are relatively low power and so they're unlikely to burn your skin or eyes. However, some communications systems do use powerful laser light sources, the light from which, or at the end of a fibre, **may** be sufficient to damage your eyesight by burning the soft tissue of the retina, cornea or lens.

Making matters worse, light used for fibre-optic cable communications is seldom in the visible spectrum so you wouldn't know whether the light source is on or off. So...

...you should **never look into a fibre or a fibre light source.**

When terminating or splicing optical fibres, you will be exposed to small glass shards from the broken/cut end(s) of the fibre. These shards can be dangerous in several ways. They may stick to your fingers then enter your eyes as you touch them causing irritation and tissue damage. They may stick to your fingers and get into your food and drink which can cause internal bleeding. They can enter your skin which can be painful in the first instance but, as glass is inert, the glass piece won't break-down and can work their way through body cutting nerves and blood vessels. Finally, the chemicals involved in fusion splicing have risks associated with them too.

Some safety rules to follow when terminating or splicing include:

- Keep food and drink away from the work area
- Always wear protective gloves and safety glasses with side shields
- Always wear a disposable apron to minimize fibre particles on your clothing
- Always wash your hands after work is complete
- Do not touch your eyes or contact lenses unless you have washed your hands
- Dispose of all fibre pieces in a safe manner
- Read the MSDS for the chemicals that you're using so you know how to store, work with, and dispose of them safely.

Student notes

Review questions

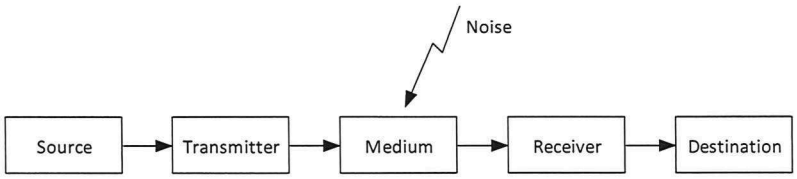
Answer these questions to check your understanding of what you have learnt from this chapter. Doing this will help you prepare for the knowledge assessment.

1. Sketch the block diagram of a basic fibre-optic communications system.
2. What type of light is normally used in fibre-optic system?
3. What material is normally used for optical fibres?
4. What is meant by the term "splice" in the context of fibre-optics?
5. List at least three advantages of optical fibre over communications using metallic conductors.
6. State the two main occupational safety hazards associated with fibre-optic work.
7. State at least five safety rules to follow when working with optical fibre.

Answers to review questions

Section 1

1. Basic block diagram of a communications system:



2. Source – Any device that produces a message signal in electrical form (so microphone, CCD, etc.)
   Transmitter – RF transmitter, fibre-optic transmitter, modem, etc.
   Medium – Free space, copper cable, fibre optic cable
   Receiver – Radio receiver, fibre-optic receiver, modem, etc.
   Destination – Loudspeaker, video screen, computer, etc.

3. These three terms are all used as names for the information to be transmitted which may be one or more of: speech, music, video, and data.

4. "Noise" is any unwanted signal that is superimposed on the desired signal. This includes things like hiss, crackle, pops etc but also the signal from another legitimate channel if it's not what's meant to be received.

5. The medium is most susceptible to noise. This is because the signal is usually at its smallest and there are many sources of noise that impart their EMR on the medium.

6. The term "modulation" refers to the process of impressing or encoding the message signal on a higher frequency carrier.

7. Modulation is used because: a) the higher frequency carrier can be transmitted more efficiently by the antenna; b) the transmit and receive antenna lengths can be much shorter; and c) it permits multiple users/channels within the same radio horizon.

8. Abbreviation definitions:

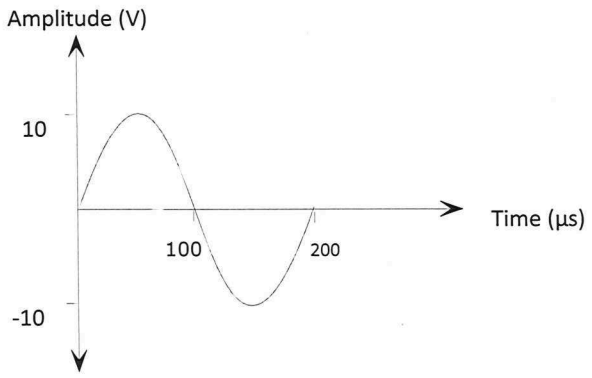
- VLF - Very low frequency
- DSBFC – Double sideband full carrier
- UHF – Ultra high frequency
- SSBSC – Single sideband suppressed carrier
- MF – Medium frequency
- FM – Frequency modulation
- VHF – Very high frequency
- PM – Phase modulation
- HF – High frequency
- AM - Amplitude modulation

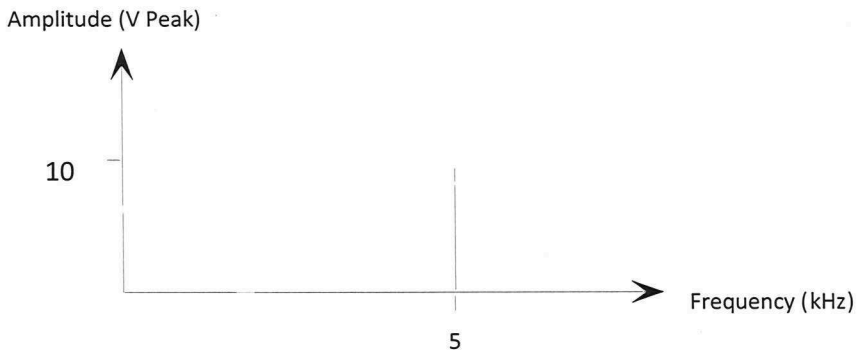
9. The channel bandwidth for AM broadcast radio is 17kHz and for FM broadcast radio it's 200kHz. For broadcast television it's 7MHz.

10. 22.73dB

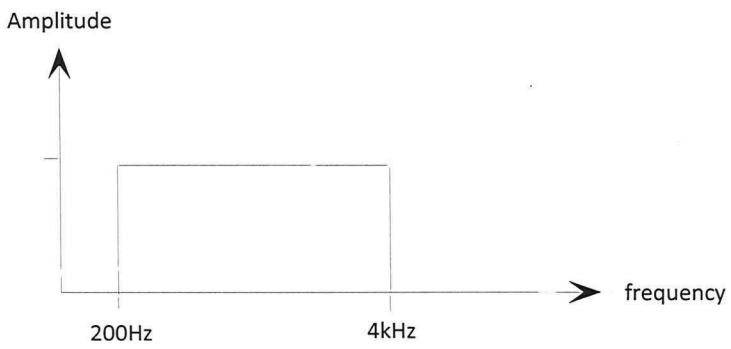
11. The SNR from power figures is **10**Log(A_P) whereas The SNR from voltage figures is **20**Log(A_V).

12. A single 5kHz sine wave tone in the time and frequency domains:





13.



Section 2

1. An electromagnetic wave consists of an electric field (E) and a magnetic field (H) that are perpendicular (ie at 90°) to each other. Both fields are perpendicular to their direction of travel (propagation).

2. Electromagnetic waves are polarised in the direction as the electric (E) field.

3. Electric and magnetic fields that are not returned to the circuit when they collapse and are instead radiated into free space.

4. Wavelength of an electromagnetic waves:

    a. 20MHz = 15m
    b. 500MHz = 0.6m

5. Ground waves (or surface waves) are radio waves that travel along the earth's surface. They must be vertically polarised.

6. Ground waves are suited to the medium frequency (MF) and lower bands.

7. Marine communications systems make use of ground waves.

8. In a "line of sight" transmission, the space between the transmitter and receiver antennas should be uninterrupted.

9. The VHF and UHF bands are best suited to space wave transmissions.

10. Reflected space waves can cause multiple out-of-phase signals to appear at the antenna in turn causing interference.

11. Sky waves are space waves that are directed towards the ionosphere where they are refracted. (They can subsequently bounce off the earth surface in a phenomenon known as "skipping" to extend the distance of sky wave communications.)

12. The ionosphere is the uppermost part of the atmosphere in which small particles have been ionised by solar radiation.

13. Longer communications distances are possible with sky waves compared to space waves because of skipping.

14. Shortwave two-way radio is an application for sky waves.

15. Geostationary satellites follow the earth's orbit and therefore appear to be stationary with respect to the earth's surface.

16. In satellite communications, the transmitting antenna sends a directional signal through the ionosphere to a geostationary satellite's receiving antenna. The satellite re-amplifies the received signals and transmits them back to earth at different carrier frequencies.

17. Satellite communications can achieve the longest transmission distances because of the height of the satellite above the earth's surface. It enables signals to be beamed back over a very wide (transoceanic) geographical area.

Section 3

1. A dipole antenna consists of two conductors mounted axially (or collinearly) to each other ideally with only a small gap between them.

2. No, the Marconi antenna is not much less efficient than the dipole. That's because, provided the Marconi antenna is mounted vertically just above the ground, or just above a counterpoise, another conductor is effectively created by reflection of the monopole in the ground/counterpoise.

3. a) Space waves; b) Ground waves.

4. An antenna is "horizontally polarised" when its E field is parallel to the earth's surface.

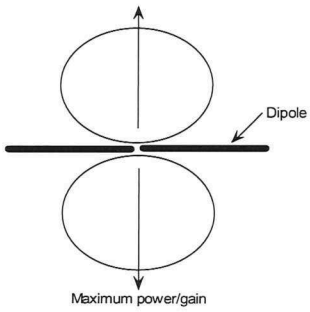
5. A receiving antenna should be in the same polarisation as the transmitting antenna so that the maximum signal is induced in the antenna.

6. Length of a half-wave dipole antenna for 182.25MHz:

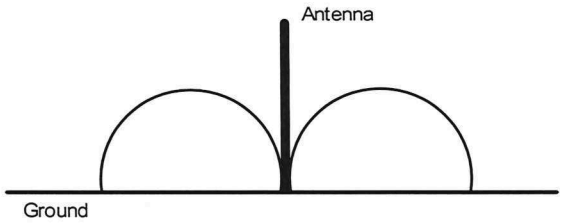
$$\lambda = \frac{v}{f} = \frac{3\times10^8}{182.25\times10^6} = 1.65\,\text{m}$$

Therefore, $\frac{\lambda}{2} = \frac{1.65}{2} = 825\,\text{mm}$

7. The radiation pattern for a dipole antenna (viewed from above):



8. The radiation pattern for a Marconi antenna (viewed from the side):



9. Calculate the length of a quarter-wave Marconi antenna at 1,269kHz.

$$\lambda = \frac{v}{f} = \frac{3\times10^8}{1269\times10^3} = 236\,\text{m}$$

Therefore, $\frac{\lambda}{4} = \frac{236}{4} = 59\,\text{m}$

10. The director and reflector in a Yagi antenna make the antenna more directional.

11. A Yagi antenna for 198.5MHz:

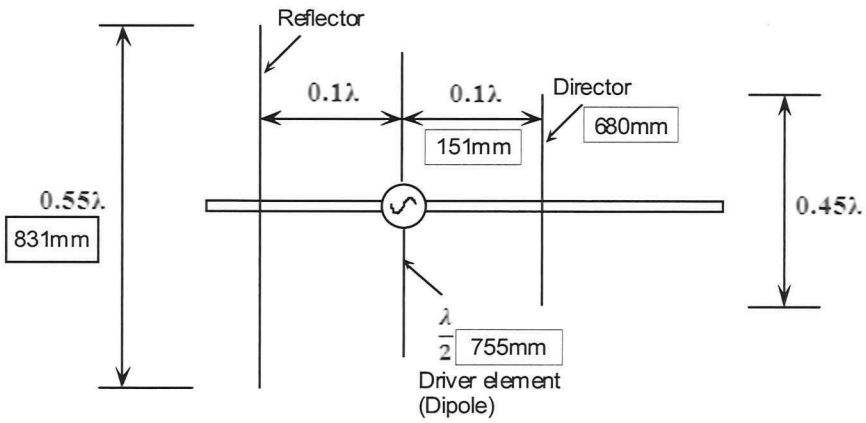
$$\lambda = \frac{v}{f} = \frac{3\times10^8}{198.5\times10^6} = 1.51\text{m}$$

Reflector length = 0.55λ = 831mm

Dipole length = 0.5λ = 755mm

Director length = 0.45λ = 680mm

Element spacing – 0.1λ = 151mm



12. The risks include:

- Electric shock from touching the antenna
- RF burns due to EM wave induced currents in your body
- RF burns of burns due to dielectric heating of body parts
- Risk of falls when working at heights

13. Electrocution when touching an antenna can occur by either: a) touching two points of the antenna with both hands and there is a sufficient potential difference between those two points; or b) touching a sufficiently high potential point of a monopole antenna while standing on the ground. In both cases, the potential is sufficient to cause fatal electrocution if it can cause a current of 20mA or more to pass across the heart/lungs.

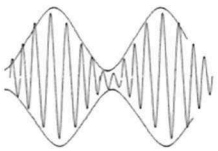
14. Refer to the notes.

Section 4

1. Amplitude modulation involves using the message to vary the carrier's peak-to-peak voltage. [Specifically, the carrier's instantaneous voltage is directly proportional to the message's instantaneous voltage.]

2. The diode or transistor.

3. A band-pass filter. This will remove all of the unnecessary sinewaves both below and above the sinewaves needed for the AM signal.

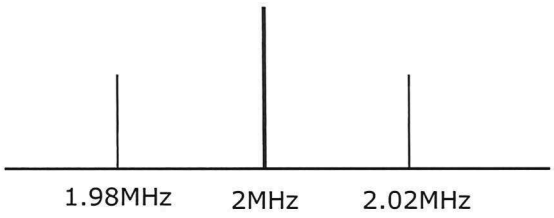
4. 

5. Double Sideband Full Carrier.

6. A 2MHz carrier amplitude-modulated by a 20kHz message signal:

    a. Sideband frequencies will be 2MHz ± 20kHz (1.98MHz and 2.02MHz)

    b. Sketch the LSB, carrier and USB in the frequency domain.



    c. BW = 2 × fm = 2 × 20kHz = 40kHz.

    d. The AM signal's bandwidth would still be 20kHz for the complex message described because the additional components in the AM signal would all be between 1.98MHz and 2.02MHz.

7. Modulation index calculations:

A. $m = \frac{B-A}{B+A} \times 100\% = \frac{2-1}{2+1} \times 100\% = 33.3\%$

B. $m = \frac{B-A}{B+A} \times 100\% = \frac{4-1}{4+1} \times 100\% = 60\%$

C. $m = 100\%$

8. It's desirable for a transmitter to use modulation levels approaching 100% to maximise the power in the sidebands (which is where the transmitted information is contained). This would increase the distance that the transmitted signal can be picked-up by a receiver.

9. Overmodulation refers to a modulation index greater than 100%. Importantly, the AM signal's upper envelope is no-longer the same shape as the message and its lower envelope is no-longer the same shape as the message inverted.

10. Over-modulation is undesirable because: 1) The resulting distorted envelopes make it impossible for the envelope detector to recover the message; and 2) Where a class C amplifier is used at the transmitter, over-modulation can cause "sideband splatter".

11. Single Sideband Suppressed Carrier.

12. An advantage of DSBFC over SSBSC is that it is easier to demodulate. A disadvantage is that that DSBFC is inefficient in that power is wasted in transmitting the carrier. Also, DSBFC requires a wider bandwidth.

13. Total power when a 20kW carrier is modulated to 88%:

$$P_t = P_c\left(1+\frac{m^2}{2}\right) = 20\times10^3 \times \left(1+\frac{0.88^2}{2}\right) = 27.7\text{kW}$$

14. The total sideband power in the previous question:

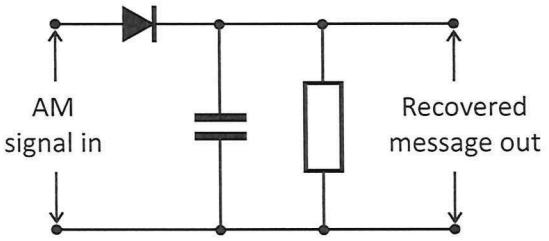
27.7kW – 20kW = 7.7kW (which is 3.85kW in each sideband).

15. Calculate the carrier power for an AM transmission that has a total power of 5kW and a modulation index of 95%?

$$5\times10^3 = P_c\left(1+\frac{0.95^2}{2}\right)$$

$$P_c = 3.45\text{kW}$$

16.



17. Advantages: Cheap (just three components), simple, produces a DC component that can be used for AGC.

18. Using the envelope detector to demodulate SSBSC signals would result in severe distortion and so the message would be unintelligible.

This is because the circuit recovers the message in DSBFC AM by tracking the signal's envelopes and this works because the envelopes of DSBFC signals are the same shape as the message. However, the envelopes of SSBSC signals are not the same shape as the message.

[An SSBSC signal requires reinsertion of the carrier signal for demodulation. Importantly, the inserted carrier must be exactly the same frequency and phase of the carrier used to generate the SSBC signal in the first instance. This makes the receiver circuitry inherently more complex than a normal AM receiver. The circuit that does this is the product detector.]

Section 5

1. In FM, the frequency of the carrier is proportional to the amplitude of the message signal.
2. A voltage-controlled oscillator's (VCO).
3. An FM carrier's deviation is the amount by which the carrier frequency increases and decreases above and below its centre (unmodulated) frequency.
4. The *rate* of deviation of an FM signal is dependent on the frequency of the message signal.
5. The amplitude of the message signal and the amount of deviation of an FM carrier are proportional. The greater the amplitude, the greater the deviation.
6. The frequency swing for a 100MHz carrier with a deviation is 50kHz is 100MHz ± 50kHz or 99.95MHz and 100.05MHz.
7. Broadcast FM uses a carrier deviation of 75kHz.
8. Modulation index for an FM system having a maximum carrier deviation of 15kHz and a message signal of 5kHz:

$$m_f = \frac{\delta}{f_i} = \frac{15kHz}{5kHz} = 3$$

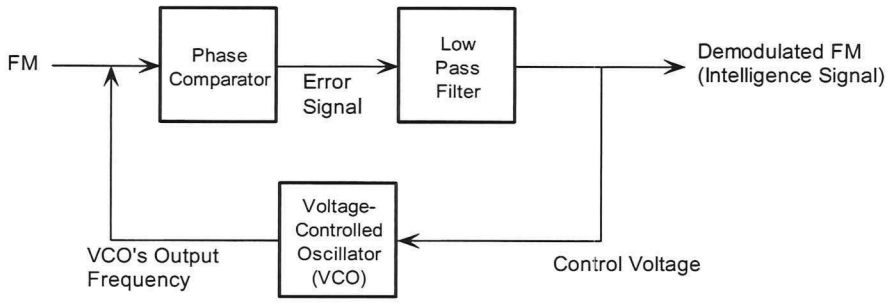
9. A table of Bessel Functions plots modulation index against relative carrier and the amplitude of pairs of significant sideband frequencies.
10. The spectral composition of the FM signal and/or the FM signal's bandwidth.
11. As the modulation index increases so do the number of significant sideband frequencies.
12. For NBFM, the BW = 2 × fm = 2 × 3kHz = 6kHz.
13. m_f = 1.5kHz/3kHz = 0.5
14. NBFM signals are those with a modulation index of 0.5 or less.
15. For WBFM, the BW = 2 × (fm + Δf) = 2 × (12kHz +75kHz) = 174kHz.
16. m_f = 75kHz/12kHz = 6.25
17. WBFM signals are those with a modulation index of 5 or more.
18. Power in an FM signal is constant because the carrier's amplitude is constant. Variations in frequency do not affect power.
19. Frequency Shift Keying (FSK) is an FM communications system in which the message is a digital signal.

Section 6

1. FM has better noise immunity compared to AM because noise in a communication system is usually manifested as an amplitude variation which is not detected by the FM detector (ideally).
2. The limiter stage in FM receiver ensures the amplitude of the FM signal is constant prior to the discriminator. (This matters if the discriminator does respond to amplitude variations in addition to frequency variations.)

   With that said, the limiter may be included in an FM receiver - regardless of whether it's needed to remove amplitude variations - to implement a form of automatic gain control (AGC). AGC is discussed in a later chapter.]
3. The purpose of pre-emphasis is to improve the signal-to-noise ratio of the high frequency components of the message signal by increasing their amplitude prior to modulation.
4. Pre-emphasis found in the transmitter prior to the modulator.
5. Pre-emphasis take effect above 3,182Hz.
6. De-emphasis negates the effect of pre-emphasis by reducing the high frequency components above 3,182Hz.
7. De-emphasis found in the receiver just after the detector.
8. A first order RC LPF.
9. The calculation for an RC LPF's cut-off frequency is: $= \frac{1}{2\pi RC}$. For the cut-off frequency to be 3,182Hz, the "R × C" component of that equation must equal 50μs.
10. An FM discriminator is simply an FM demodulator (also called an FM detector). It recovers the message signal from the FM signal. (That said, your author hasn't ever read the term "discriminator" used for the PLL.)
11. An ideal FM discriminator produces an output signal that is directly proportional to the frequency deviation of the carrier.
12. Slope detector, Foster-Seeley discriminator, ratio detector and quadrature detector.

13. Block diagram and basic operation of a phase-locked loop (PLL) FM detector:



The PLL uses a phase comparator to compare the phase (and hence the frequency) of the incoming FM signal with that of a local voltage controlled oscillator (VCO). The comparator produces an error signal at its output that it proportional to the difference between its two input-signals. This error signal is low-pass filtered and used as the control voltage for the VCO. The input to the VCO (out of the low-pass filter) is a copy of the message signal.

14. (These days) The PLL can be wholly implemented inside an integrated circuit (IC).

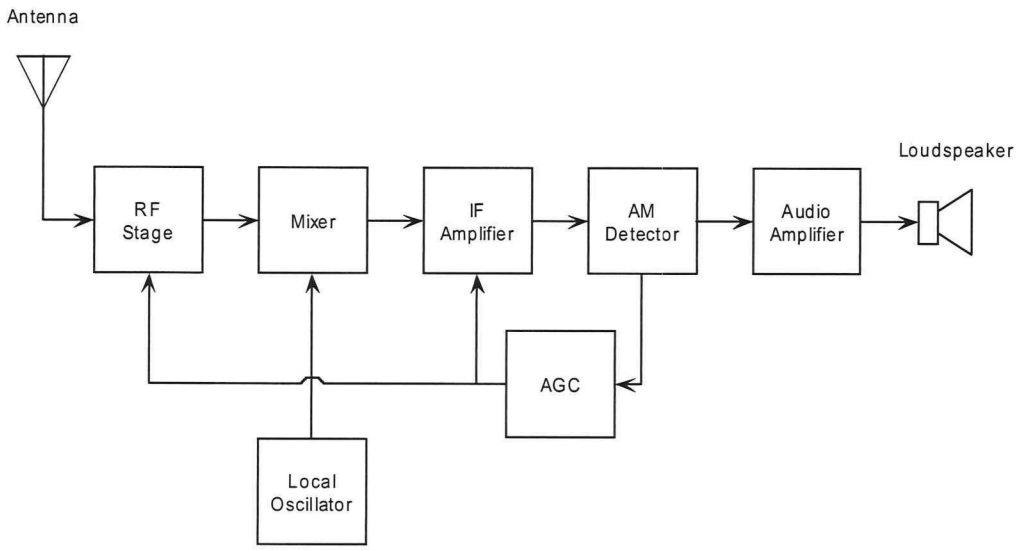
This is important because Foster-Seeley discriminator and Ratio detector need a transformer which are large and cannot be implemented inside a transformer. A better option (before the PLL became widely) used was the Quadrature detector. It doesn't need a transformer but it still needs an external LC circuit which means it cannot be implemented inside an IC but at least it was smaller than a transformer.

Section 7

1. The sensitivity of a radio receiver is the input signal voltage at the antenna required to produce a specific output voltage ("output" here is normally the output of the demodulator).
2. The selectivity of a radio receiver describes "how well" it is capable of differentiating between the required signal (channel) and all other unwanted signals such as adjacent channels.
3. A low or poor selectivity would result in adjacent channels being audible. The noise (hiss, crackles, etc) would likely be greater too.
4. An overly high selectivity could mean that some of the sidebands, that contain the message signal, do not get through, resulting in loss of audio fidelity (quality).
5. The requirements for a basic receiver include: an antenna (ideally efficient), RF tuning circuit, demodulation, audio amplification, and a loudspeaker.
6. The main limitation of the TRF receiver was the problem associated with its selectivity which was not consistent across the band that it's designed to tune across.

Section 8

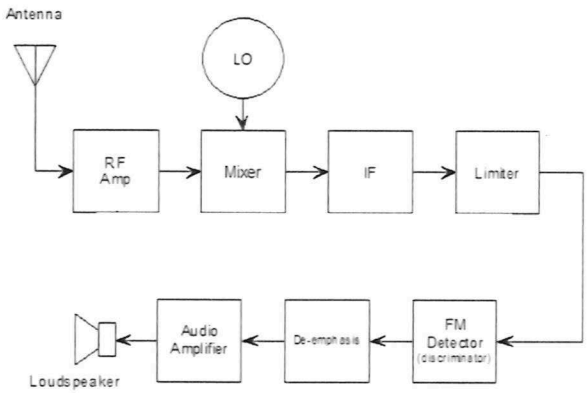
1. In a superheterodyne receiver, the received channel is down-converted to a lower frequency called the intermediate frequency (IF) before being demodulated.
2. Block diagram for an AM superheterodyne receiver:



3. Function of the following blocks:
   - Mixer - Non-linearly "mixes" the AM signal and the local oscillator to produce a set of sum and difference components.
   - Local oscillator – Is tuned to be higher than the received signal by the IF frequency. This causes one of the down-converted carrier at the output of the mixer to be at the IF.
   - IF amplifier – One or several of band-pass amplifier stages that provide the majority of the receiver's sensitivity and all of its selectivity.
   - AGC - *Automatic Gain Control* - adjusts the gain of the receiver depending on the strength of the received signal so that the sound out is constant (for a particular volume setting.

4. The centre frequency of the IF amplifier in an AM receiver is 455kHz.
5. The centre frequency of the IF amplifier in an FM receiver is 10.7MHz.
6. A superheterodyne receiver has superior selectivity compared to a TRF receiver because each channel is down-converted to a common IF. As the IF stage's centre frequency is fixed, it's a relatively simple matter to set its Q for the appropriate bandwidth. (Actually, further superior selectivity is achievable using ceramic or crystal filters.)
7. The local oscillator frequency in an AM superheterodyne receiver tuned to 1134kHz will be 1134kHz + 455kHz = 1589kHz (or 1.589MHz).
8. 455kHz, 450kHz & 460kHz **and** 2723kHz, 2718kHz & 2728kHz.
9. An image frequency is any unwanted frequency that causes one of the components at the output of the mixer to be at the IF frequency, and therefore pass through the IF stage.
10. An image frequency can get through the IF stage when it is higher (or lower) than the wanted frequency by twice the IF (or other multiples).
11. Image = 1134kHz + (2 × 455kHz) = 2044kHz.
12. Although the RF Stage is normally a BPF with a wide bandwidth, it's not so wide as to let the image through to the Mixer'.
13. The automatic gain control (AGC) adjusts the gain of the receiver depending on the strength of the received signal to give a constant signal level at the output of the detector. Without AGC, the user would always have to turn the volume up for weak signals and down for strong signals.
14. The AGC control voltage is effectively the DC component on the output of the envelope detector. This voltage is fed back to the IF and/or RF stages which use it to vary the gain of the stage and affect the sensitivity of the receiver.

15.



16. The purpose of the limiter in an FM superheterodyne receiver is to provide a constant amplitude output. This removes any unwanted noise or amplitude modulation from the FM signal to provide a constant amplitude signal to the FM detector (discriminator).

17. AGC is not necessary in FM superheterodyne receivers with a limiter because the limiter provides a constant input voltage to the FM detector.

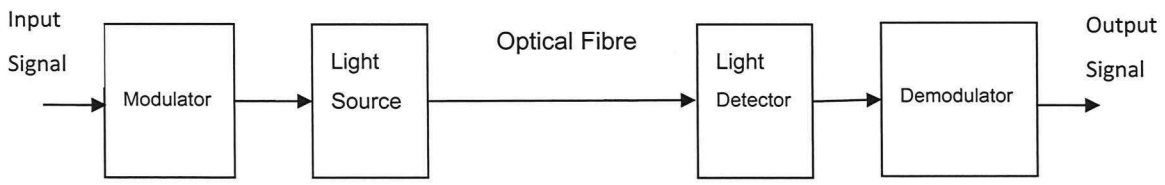
18. The typical frequency range of the local oscillator in a broadcast FM receiver (assuming the FM band to be 88.1MHz to 107.9MHz) would be 98.8MHz (ie 88.1MHz + 10.7MHz) to 118.6MHz (ie 108MHz + 10.7MHz).

19. 102MHz

20. 112.7MHz

Section 9

1. Sketch the block diagram of a basic fibre optic cable communications system.



(**Note:** You may notice that this is basically the block diagram of a ommunications system on page 8.)

2. Infrared light is normally used in fibre-optic systems. Visible light is sometimes used for low-cost plastic fibre systems.
3. Optical fibres are normally made of glass. Low cost systems are made from plastic.
4. The term "splice" in fibre-optics means to join without using a connector.
5. At least five advantages of optical fibre over copper cable:
   - Much lower signal attenuation
   - Extremely wide bandwidths permitting extremely high data rates
   - Immunity to electromagnetic interference (EMI)
   - No crosstalk
   - More secure point-to-point connection

6. Two main occupational safety hazards associated with fibre-optic work include:
   - Light sources causing possible damage to the naked eye
   - Tiny scraps of glass fibre can get into the eyes or be ingested.

7. At least five safety rules to follow when working with optical fibre include:
   - Never look directly into the end of fibre cables. A fibre optic power meter can be used to make certain there is no light being transmitted.
   - Keep food and drink away from the work area as the particles can be ingested which can cause internal haemorrhaging.
   - Always wear safety glasses with side shields and protective gloves.
   - Wear disposable aprons to minimize fibre particles on your clothing.
   - Always wash your hands.
   - Do not touch your eyes or contact lenses unless you have washed your hands.
   - Dispose of all fibre pieces in a safe place.
   - Thoroughly clean your work area when you are done

Textbooks, eBooks and websites

Textbooks

G.M. Miller (2007). *Modern electronic communication* (9th edition). Prentice-Hall.

W. Tomasi (2003). *Electronic communications systems* (5th edition). Prentice Hall.

L. Frenzel (2015). *Principles of electronic communications systems* (4th edition). McGraw-Hill.

The ARRL Radio Amateurs Handbook. (Annual publication)

Paul H. Young (2003). *Electronic Communication Techniques* (5th edition). Pearson.

W. Schweber (2001). Electronic *Communication Systems: A Complete Course*, (4th edition). Prentice-Hall.

eBooks

Newnes Radio and RF Engineering Pocket Book By: Winder, Steve; Carr, Joseph; Davies, John. Published By: Elsevier Edition: 3
http://www.ebooks.com/ebooks/book_display.asp?IID=305679

Websites

List of Australian radio stations:

http://www.ausradiostations.com/

List of Australian TV channel frequencies:

http://www.transmissionservices.com.au/Analogue%20&%20Digital%20TV%20Channel%20Frequencies.pdf

Selected experiments for EC0063 from the Emona 101 Trainer Lab Manual

From Volumes 1, 2 & 3
Experiments in Modern Analog & Digital Telecommunications
For use with the PicoScope

Barry Duncan & Carlo Manfredini

analog
digital
biskit™

Selecteed experiments for EC0063 from the Emona-101 Trainer Lab Manual Volumes 1, 2 & 3
Experiments in Modern Analog and Digital Telecommunications.

Authors: Barry Duncan and Carlo Manfredini

Issue Number:

Published by:

Emona Instruments Pty Ltd,
78 Parramatta Road
Camperdown NSW 2050
AUSTRALIA.

web: www.ett101.com
telephone: +61-2-9519-3933
fax: +61-2-9550-1378

Copyright © 2006 - 2021 Emona Instruments Pty Ltd and its related entities. All rights reserved. No part of this publication may be reproduced or distributed in any form or by any means, including any network or Web distribution or broadcast for distance learning, or stored in any database or in any network retrieval system, without the prior written consent of Emona Instruments Pty Ltd.

For licensing information, please contact Emona Instruments Pty Ltd.

The "biskit" logo is a trademark of Emona TIMS Pty Ltd

Printed in Australia

Introduction

The ETT-101 Trainer

As its name implies, the Emona Telecoms-Trainer 101 is used to help students learn about communications and telecommunications principles. It lets you bring to life the block diagrams that fill communications textbooks. A "block diagram" is a simplified representation of a more complex circuit. An example is shown in Figure 1 below.

Block diagrams are used to explain the principle of operation of electronic systems (like a radio transmitter for example) without worrying about how the circuit works. Each block represents a part of the circuit that performs a separate task and is named according to what it does. Examples of common blocks in communications equipment include the *adder, multiplier, oscillator*, and so on. *oscillator*, and so on.

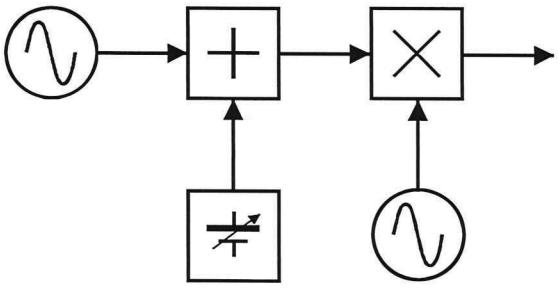
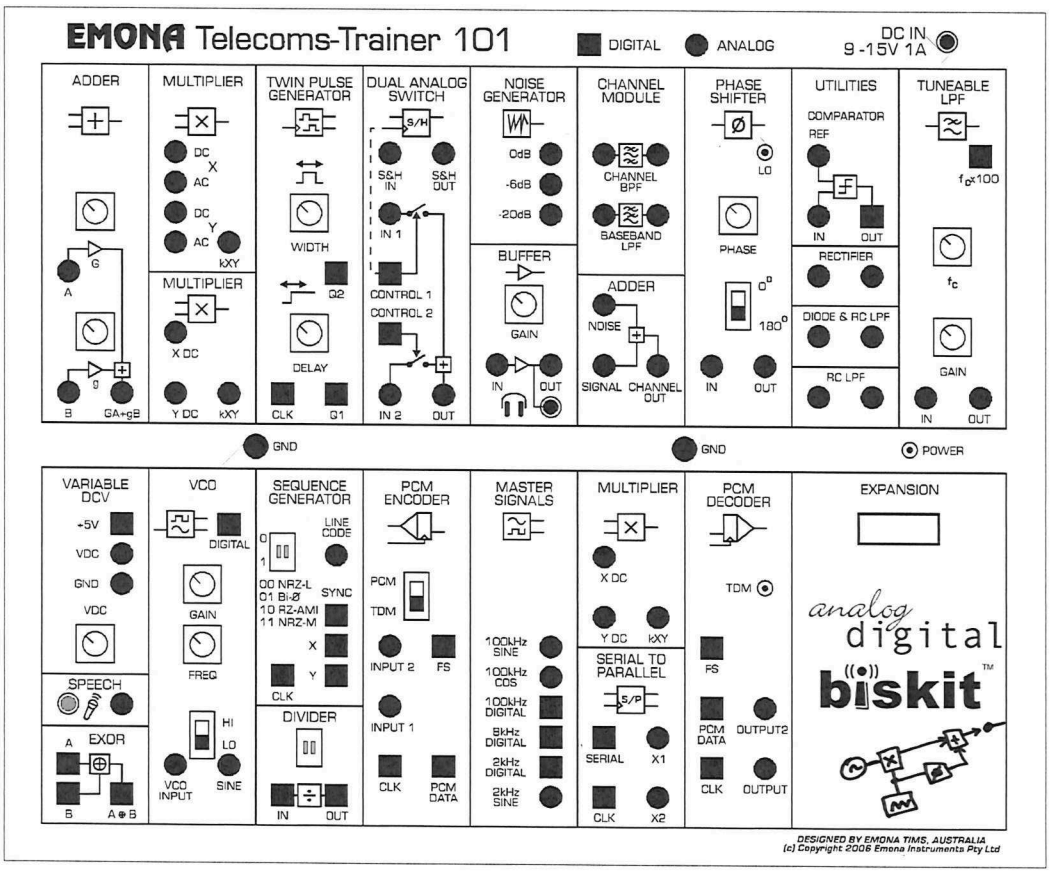


Figure 1

The Emona Telecoms-Trainer 101, illustrated below, has a collection of blocks (called *modules*) that you can put together to implement dozens of communications and telecommunications block diagrams.



The ETT-101 Lab Manual

The ETT-101 Lab Manual Volume One, is written specifically for students in the senior years of technical high schools and vocational colleges. The experiments are intended to support the theoretical concepts introduced in class work. It is assumed that the students only have a basic knowledge of mathematics and a limited background in physics and electricity.

The 19 chapters cover a broad range of telecommunications concepts, from fundamental topics familiar to all students, such as AM and FM broadcasting, through to the latest technologies used in mobile telepohnes and wireless networking. In each experiment, the core technology is revealed to the student, at its most fundamental and simplest level.

In order to make the student's learning experience more memorable, the student is usually able to both view signals on the oscilloscope and then listen to their own voice undergoing the modulation or coding scheme being investigated.

Another important factor which makes the learning experimence more valuable for the student is the ETT-101 Trainer's experimental approach: the student is "allowed to make mistakes". As the student builds the experiment, they need to make constant observations and adjustments. If signals are not as expected then the student needs to make a decision as to whether the correction required is an adjustment or an incorrect interconnection.

Each experiment in the ETT-101 Lab Manual provides a basic introduction to the topic under investigation, followed by a series of carefully graded hands-on activities. At the conclusion of each sub section the student is asked to answer questions to confirm their understanding of the work before proceeding.

Chapters can be covered in any order, however, it is imperative that all students complete the first three before proceeding to the subsequent chapters. Chapter 1 introduces the use of an oscilloscope to view electrical signals. Chapter 2 follows with an introduction to the functions of the ETT-101 and chapter 3 introduces the concept of mathematical modeling using electronic functional blocks.

Since the ETT-101 Trainer is a true modeling system, the instructor has the freedom to modify existing experiments or even create completely new experiments to convey new concepts to students.

Finally, it should be noted that the ETT-101 Trainer can implement many more experiments than are documented in this Volume One Lab Manual and further experiments can be found in later manuals.

Name: _____

Class: _____

2 - An introduction to the Telecoms-Trainer 101

Experiment 2 - An introduction to the Telecoms-Trainer 101

Preliminary discussion

As its name implies, the Emona Telecoms-Trainer 101 is used to help people learn about communications and telecommunications principles. It lets you bring to life the block diagrams that fill communications textbooks. A "block diagram" is a simplified representation of a more complex circuit. An example is shown in Figure 1 below.

Block diagrams are used to explain the principle of operation of electronic systems (like a radio transmitter for example) without worrying about how the circuit works. Each block represents a part of the circuit that performs a separate task and is named according to what it does. Examples of common blocks in communications equipment include the *adder, filter, phase shifter* and so on.

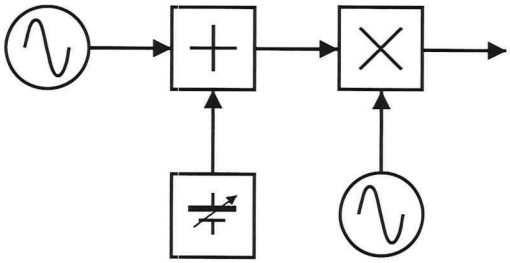


Figure 1

The Emona Telecoms-Trainer 101 has a collection of blocks (called *modules*) that you can put together to implement dozens of communications and telecommunications block diagrams.

The experiment

This experiment is in three stand-alone parts (2-1, 2-2 and 2-3) and each introduces you to one or more of the trainer's analog modules.

It should take you about 1 hour to complete 2.1 of this experiment, about 50 minutes to complete 2.2 and about 1 hour to complete 2.3.

Equipment

- Emona Telecoms-Trainer 101 (plus power-pack)
- Desktop/laptop with PicoScope 6 installed
- two Emona Telecoms-Trainer 101 oscilloscope leads
- assorted Emona Telecoms-Trainer 101 patch leads
- **For 2.1 only** - one set of headphones (stereo)

Some things you need to know for the experiment

This box contains definitions for some electrical terms used in this experiment. Although you've probably seen them before, it's worth taking a minute to read them to check your understanding.

Two signals that are **in phase** with each other are synchronised. That is, they go up and down at the same time.

Two signals that **out of phase** are not synchronised. That is, they are out of step with each other. An example of two signals that are out of phase is shown in Figure 3 below.

Phase difference describes how much two signals are out of phase and is measured in degrees (like degrees in a circle). Signals that are in phase have a phase difference of 0°. Signals that are out of phase have a phase difference > 0° but < 360°.

A **sinewave** is a repetitive signal with the shape shown in Figure 1.

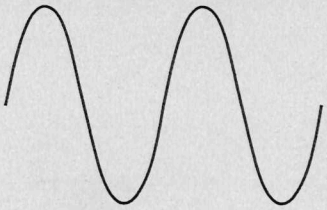


Figure 2

A **cosine wave** is simply a sinewave that is out of phase with another sinewave by exactly 90°. A sinewave and a cosine wave are shown in Figure 3. (They're not marked because, in this case, it doesn't matter which one is which.)

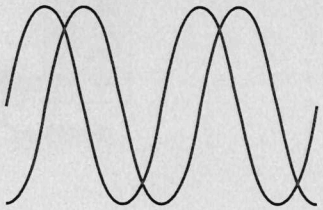


Figure 3

2.1 - The Master Signals, Speech and Buffer modules

The Master Signals module

The Master Signals module is an AC signal generator or *oscillator*. The module has six outputs providing the following:

Analog

- A 2kHz sinewave
- A 100kHz sinewave
- A 100kHz cosine wave

Digital

- A 2.083kHz squarewave
- An 8.33kHz squarewave
- A 100kHz squarewave

Each signal is available on a socket on the module's faceplate that's labelled accordingly.

Procedure

1. Gather a set of the equipment listed on page 2-2.

2. Connect the set-up shown in Figure 1 below.

   Note: Use the PicoScope for the oscilloscope pictured.

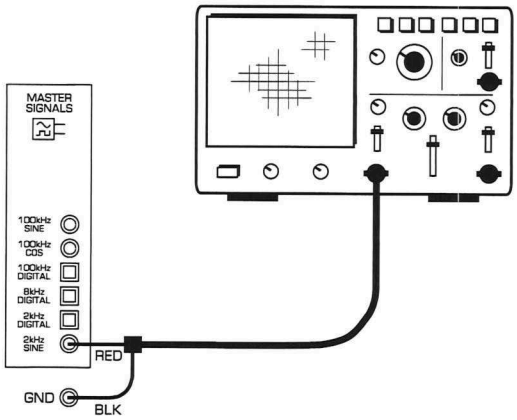


Figure 1

This set-up can be represented by the block diagram in Figure 2 below.

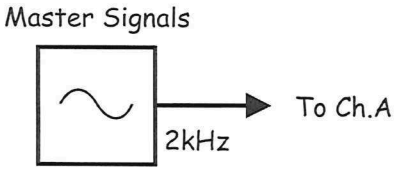
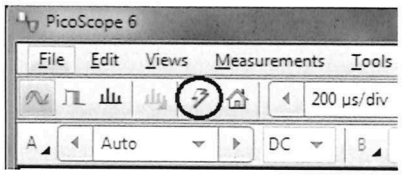


Figure 2

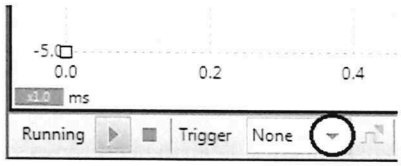
Using PicoScope 6 - Setting up

This experiment includes these breakout boxes to help you get started with using the Picoscope and to help you use some of its many features. Follow these instructions closely.

a) Launch PicoScope 6 and locate the scope's Auto Setup button (it's at the top left of the app and is circled in the adjacent screenshot). Press it once.



b) Locate the scope's Trigger Mode menu window (it's at the bottom left of the app and is circled in the adjacent). Click on it then select "Auto" from the pop-up list.



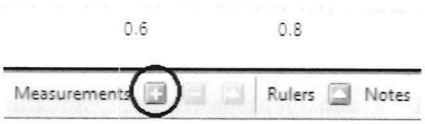
3. Use the scope's grid together with the scale information on its Y-axis to determine the peak-to-peak voltage of the Master Signals module's *2kHz SINE* output.

   Note: You should obtain a figure very close to 4Vpp. If not, call the instructor for help on how to do this.

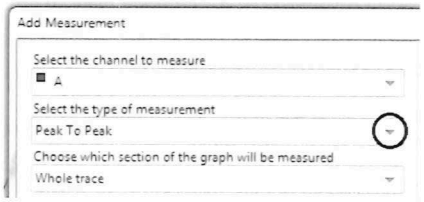
Using PicoScope 6 - Adding measurements

Like digital scopes, the PicoScope lets you measure key attributes of signals without having to rely on the display's grid. The following instructions show you how.

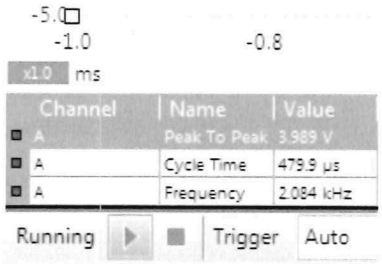
a) Locate the scope's Add Measurements button (bottom right of app) and press it once.



b) In the new window that opens, click on the "Select the type of measurement" options window and choose "Peak to Peak" from the drop-down list. Leave the other two options on their default values and press the window's OK button. This measurement will appear at the bottom of the scope's display.



c) Repeat the previous step to add measurements of the signal's period (called "cycle time" in PicoScope) and frequency. Once done, your measurements should look like the adjacent screenshot.



Notice that some measurements change value a few times per second. There are multiple reasons for this but the variation between values will be very small and can be ignored here. If you're struggling to see a clear number, pause the scope's data acquisition by pressing on the stop button (bottom left of app) which freezes the measurements. Keep in mind though that this will also freeze the scope's main display so you'll need to restart the data acquisition so that the display will update when you make changes.

Notice also that each measurement displays additional information including the minimum, maximum and average values as well the standard deviation and capture count. This information can be ignored when using the scope for this exercise.

4. Use these measurements to obtain and record in Table 1 (below) the peak-to-peak voltage, period and exact frequency of the Master Signals module's *2kHz SINE* output.

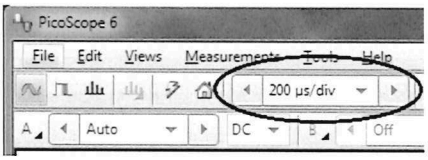
5. Connect the scope's Channel A to the Master Signals module's *100kHz COS* output.

   Note: There are now far too many cycles in the signal to see each one individually but the scope can be adjusted.

Using PicoScope 6 - Adjusting the timebase

To change the number of cycles of a signal on the scope's display, you'll need to adjust its horizontal scale (or timebase) like you would on a regular scope.

a) Locate the scope's timebase control (top left of app) and set it to 5µs/div.



6. Repeat step 4 for the Master Signals module's *100kHz COS* output.

7. Connect the scope's Channel A to the Master Signals module's *100kHz SINE* output then repeat step 4 again.

Table 1	Output voltage	Period	Frequency
2kHz SINE	3.7 V p-p	480µS	2.09 KHz
100kHz COS	4.07 V p-p	10µS	100 KHz
100kHz SINE	3.8V p-p	10µS	100 KHz

Ask the instructor to check your work before continuing.

When learning about communications theory, it's necessary to consider representations of signals in both the time domain and the frequency domain. Time domain representations show how the signal varies over time and this is what you see on the display of an oscilloscope (recall that the x-axis of a scope's display is time).

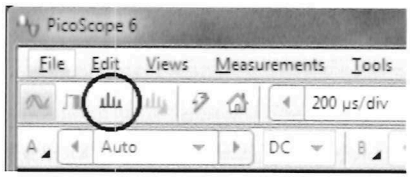
Frequency domain representations of signals plot individual sinewaves that make up the signal (it's "spectral composition") and this is what you see on the display of a spectrum analyser. Each sinewave in the signal is a vertical line projecting up from the x-axis with their position being proportional to their frequency.

The PicoScope software allows you to view signals in the frequency domain as well as the time domain.

Using PicoScope 6 - Using Spectrum Mode

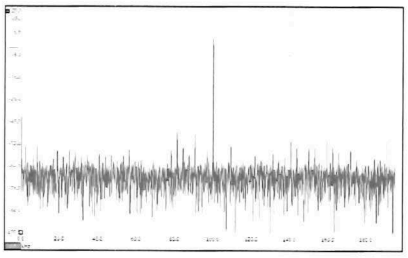
To activate the PicoScope's spectrum mode...

a) Locate the scope's Spectrum Mode button (top left of app) and press it once.



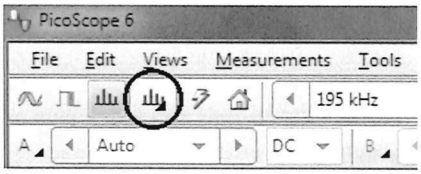
b) Set the analyser's Spectrum Range (top left of app) to 195kHz.

The display is now showing you the composition of the spectrum for all frequencies between 0Hz and 195kHz on the Channel A input. It should look like the display in the adjacent screenshot.



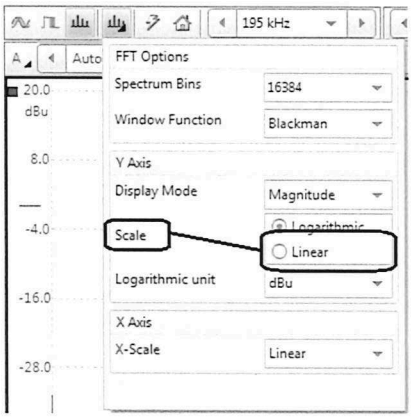
The Master Signals module's 100kHz SINE output is the large vertical line near the centre and lines up with the 100kHz vertical line of the grid. The other vertical lines on the display are sinewaves in noise that the coaxial lead connected to the Channel A input is picking up and possibly some harmonics present in the 100kHz sinewave (no sinewave is pure). These other sinewaves look large but, in fact, they're very small (the largest is about 1/350th the size of the 100kHz sinewave. They only look as large as they do because the vertical axis is logarithmic. More on this later.

b) To see the true relative sizes of the sinewaves in the spectrum, locate the analyser's Spectrum Options button (top left of app) and press it once.

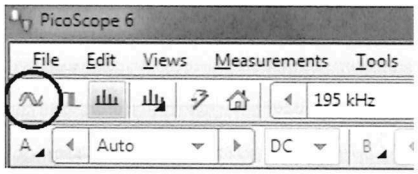


c) In the options window that opens, select Linear for the Y-axis scale then click anywhere on the display to close the window.

The Master Signals module's 100kHz sinewave should be the only spectral component that's now visible which gives you a good indication of how much smaller these other sinewaves must be.



d) To turn the Spectrum Mode off, locate the Scope Mode button (top left of app) and press it once.



It's often the case in communications that we want to view two signals at once, usually to compare them. The next step gets you to do that with two signals that are the same frequency but out of phase by a fixed amount.

8. Connect the set-up shown in Figure 3 below.

   Note: Ensure that the black plug of at least one of the scope's leads is connected to ground.

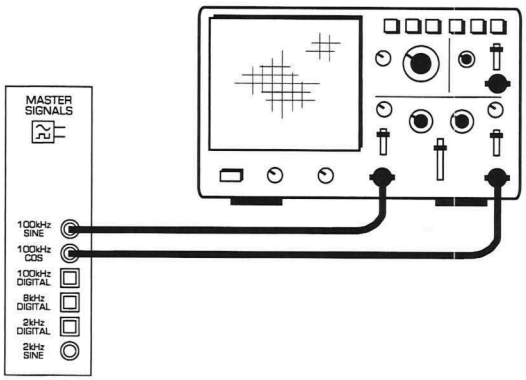
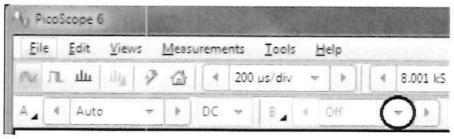


Figure 3

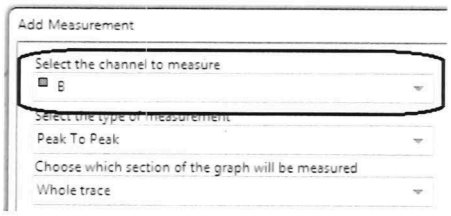
Using PicoScope 6 - Displaying both channels

To display the signals on both of the scope's channels at the same time and measure their attributes...

a) Locate the scope's Channel B vertical control (top left of app) and set it to "Auto". This turns on the Channel B and the signal on the signal on its input will be displayed.

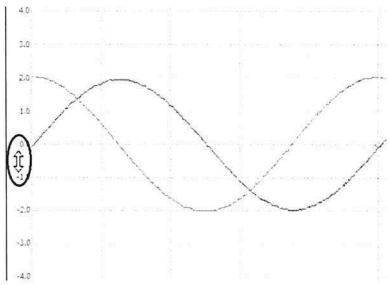


b) Use the procedure in the Adding Measurements breakout box (page 2-6) to add measurements of peak-to-peak voltage, period and frequency for the signal on Channel B as well as Channel A. When adding the measurements, you'll need to choose Channel B as shown in the adjacent screenshot.



Viewing two signals that are overlayed can be inconvenient. Adjusting the vertical position of signals displayed on a PC-based scope is quite different to regular scopes. The following instructions explain how it's done.

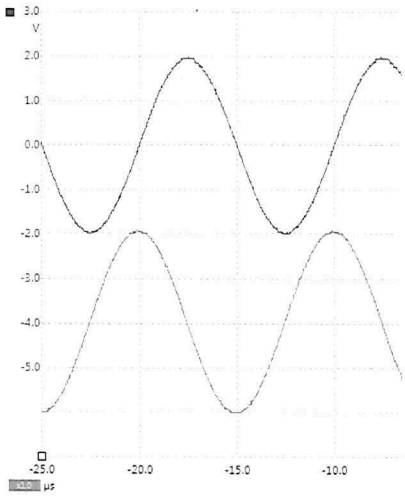
a) Move the mouse's cursor to the left side of the display's Y-axis. Once the cursor is off the grid, it'll change from a single-ended arrow to a double-ended arrow (see adjacent screenshot).



b) Click-and-hold the mouse's left button and move the cursor up and down. The Channel A signal will move with it. Reposition Channel A so that it's in the top half of the display.

c) To move the Channel B signal do the same thing but on the right side of the grid. Reposition Channel B so that its' in the bottom half of the display.

(As an aside, the vertical position of the plot on the spectrum analyser's display can be moved up and down this way too.)



It is critical to the operation of several communications and telecommunications systems that there be two (or more) sinewaves that are the same frequency but out of phase with each other. Figure 4 below shows examples of pairs of sinewaves that are out of phase by different amounts.

Figure 4

It's a useful skill in communications and telecommunications to be able to measure the phase difference between signals. The next part of the experiment gives lets you practise this by getting you to measure the phase difference between the Master Signals module's *100kHz SINE* and *COSINE* outputs.

9. Adjust the scope's *Timebase* control so that no more than two **complete** cycles of the two signals are displayed (use 2μs/div for our purposes here).

10. Set the input coupling for both channels to *AC* (top left next of the app as indicated in Figure 5 below).

11. Adjust the vertical positions of the two signals so that where they cross 0V on the same horizontal line on the grid as each other as shown in Figure 5 below.

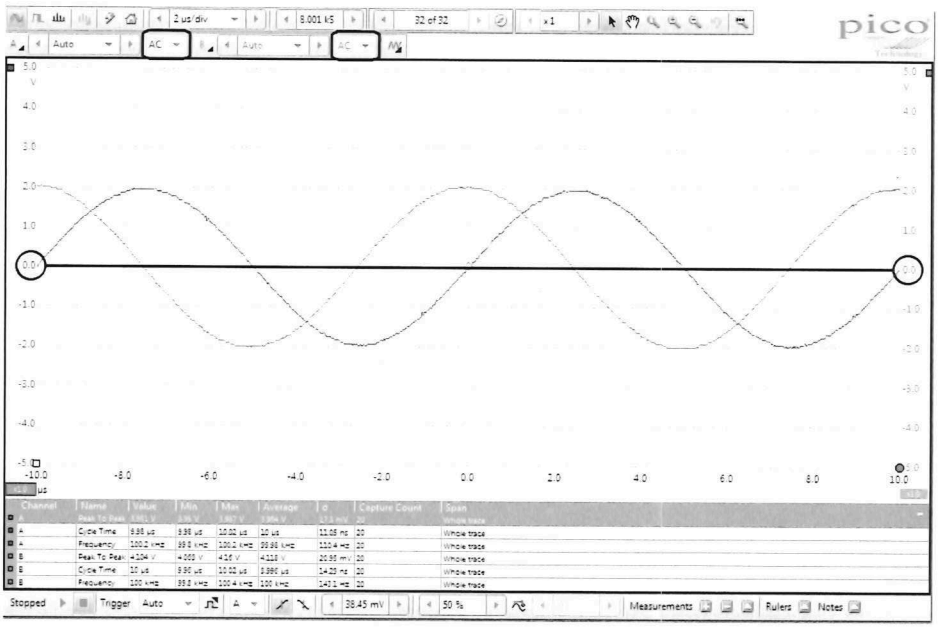


Figure 5

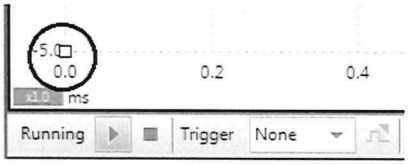
12. Use the PicoScope's measurements to determine the period of either one of the signals (their periods are the same). Recall that this is the "cycle time" measurement.

Record your measurement here: Period = 10 μS

Using PicoScope 6 - Using rulers

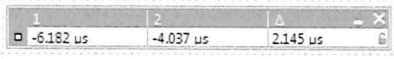
For the next step, it's useful to know how to use the PicoScope's rulers.

a) Locate the rulers' markers at the bottom left of the scope's display (there are two of them, one on top of the other).



b) Click-and-hold the first ruler's marker and drag it to any point across the screen. When you do, a dialogue box opens (see adjacent screenshot for an example). This box can be moved if it's not in a convenient location.

c) Repeat the previous step for the second ruler's marker and you'll notice that the rulers dialogue box updates with information in all three cells.



The rulers dialogue box shows three values. Reading from the left, the first value is the first ruler's position on the horizontal axis relative to the trigger point (the little yellow diamond). If the ruler is to the left of the diamond, the time value is negative to indicate that it's aligned with a point in the waveform that occurred before the trigger point. The second value is the second ruler's position on the horizontal axis relative to the trigger point. Importantly, the third value is the difference between the two rulers which is very useful for measuring the time between events on screen. (As an aside, similar ruler options are also available on the spectrum analyser's display.)

13. Use the rulers to measure the time difference between the two signals as shown in Figure 6.

Call this dimension the "difference" and record your measurement here:

Difference = 2.5 µS

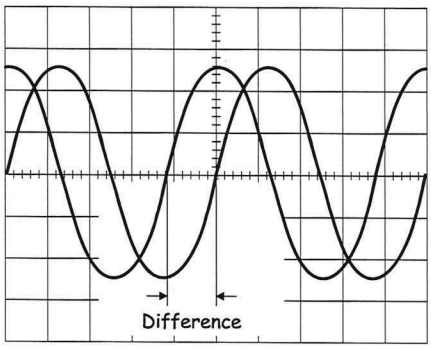


Figure 6

14. Use the equation below to calculate the phase difference between the two signals. Record the answer to your calculation in Table 2 below.

$$\Phi\ difference = \frac{difference}{period} \times 360°$$

Table 2

Φ difference between the SINE and COS outputs	90.2°

Question 1

The theoretical phase shift between a sinewave and a cosine wave is 90°. If your measurement wasn't exactly 90°, what could explain the difference?

- Noise / distortion
- Component tolerances

Ask the instructor to check your work before continuing.

15. Close the rulers' dialogue box.

    Note: This will return the rulers to the left side of the display.

16. Return the input coupling for both channels to *DC*.

17. Turn off Channel B for now.

18. Modify the set-up as shown in Figure 7 below.

    Note: Just leave the scope's Channel B connection where it is for now, you're not using Channel B so there's no need to remove it.

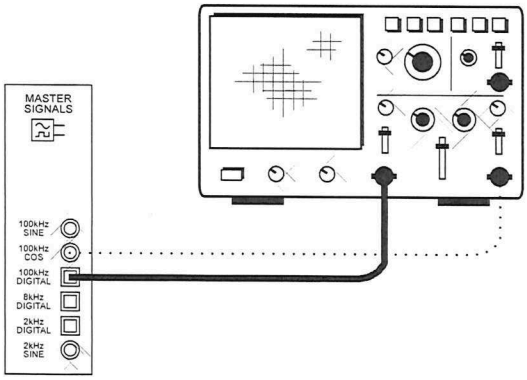


Figure 7

Notice that the digital signal is not stable on the scope's display, it moves around little on the horizontal axis. This means that the scope isn't "triggered".

You didn't have this problem when the signal connected to Channel A was the *100kHz SINE* and *COS* output because setting the Trigger's mode to *Auto* sets the trigger level to 0V by default and the sinewave is bipolar (that is, its peaks straddle 0V). However, the Master Signals module's *100kHz DIGITAL* output is unipolar (being a digital signal) so there are a theoretically infinite number of points in the waveform that is 0V. Whenever the display is updated, the scope has no way of preferring any one of those points over the others and so the point that is chosen each time is fairly random causing the waveform to move left and right. This is easily fixed by making simple adjustments to the scope's trigger settings (see next breakout box).

Using PicoScope 6 - Simple trigger adjustments

Some adjustments to the scope's trigger control can be necessary to obtain a stable signal on the scope's display. Sometimes you have a stable display but adjustments are needed make the trigger point more convenient, particularly when comparing waveforms. These instructions will show you some of the more simple controls.

a) Locate the trigger level window and type 2.5 into it to set the trigger point to 2.5V. This will give a stable display. Note the position of the yellow trigger diamond now.

b) Locate the scope's two triggering edge buttons and switch back and forth between positive and negative edge triggering while watching the signal on the display. Return to positive edge triggering when you're done.

Being complex waveforms, square waves are made of a multiple sinewaves. Recall that to render the squarewave's shape, the lowest frequency sinewave (called the fundamental) has the same frequency as the squarewave itself with all of the other sinewaves being odd numbered multiples of the fundamental frequency. Recall also that the sizes of these sinwaves must follow a pattern with the fundamental being approximately the same size as the squrewave, the third harmonic being one third the size of the squarewave and so on.

19. Activate the Spectrum Mode to view the spectral composition of the Master Signals module's *100kHz DIGITAL* output.

20. Set the analyser's input coupling for Channel A to *AC*.

21. Set the Spectrum Range to *2MHz*.

22. Set the Y-axis Scale to *Linear* (if it's not already).

23. Use a horizontal ruler to verify the frequency of the signal's fundamental and any significant harmonics that you can see.

24. Use a vertical ruler (blue marker at the top left corner of the analyser's display) to note the peak voltage of the signal's fundamental and the significant harmonics.

Ask the instructor to check your work before continuing.

The Speech module

Sinewaves are important to communications. They're used extensively for the *carrier* signal in many communications systems. Sinewaves also make excellent test signals. However, one of the key purposes of communications equipment is the transmission of speech and so it's useful to examine the operation of equipment using signals generated by speech instead of sinewaves. The Emona Telecoms-Trainer 101 allows you to do this using the Speech module and you will investigate its operation next.

25. Return to Scope Mode.

26. Return the scope's trigger level to 0V.

27. Set the scope's Channel A vertical scale to *±20V* instead of *Auto*.

28. Set the scope's timebase to *5ms/div*.

29. Connect the set-up shown in Figure 8 below.

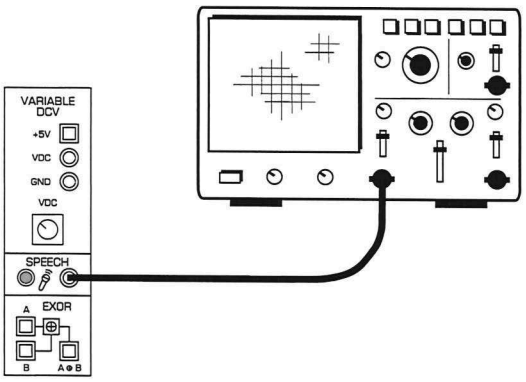


Figure 8

30. Talk into the microphone while watching the scope's display. Be sure to say "one" and "two" several times. Try humming and whistling too.

31. Turn your mobile phone's speaker on and play a song. Turn the volume down so you can barely hear the song playing then hold the phone's speaker close to the Speech module's microphone.

Ask the instructor to check your work before continuing.

The Buffer module

Amplifiers are used extensively in communications and telecommunications equipment. They're often used to make signals bigger. They're also used as an interface between devices and circuits that can't normally be connected. The Buffer module in the Emona Telecoms-Trainer 101 is an amplifier that can do both.

32. Locate the Buffer module and set its *Gain* control to about the middle of its travel.

33. Connect the set-up shown in Figure 9 below.

    Note: Ensure that the black plug of at least one of the scope's leads is connected to ground.

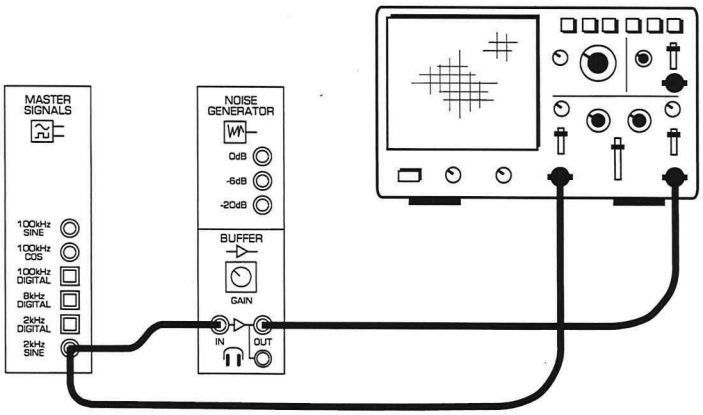


Figure 9

This set-up can be represented by the block diagram in Figure 10 below.

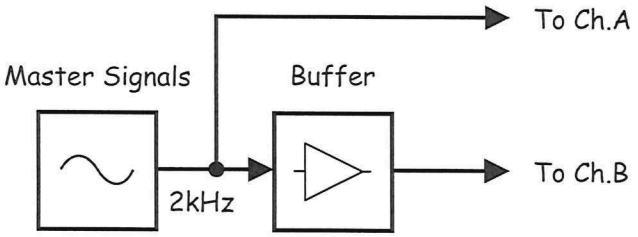


Figure 10

34. Activate the scope's Auto Setup.

35. Turn Channel B back on.

36. Set the vertical scale of Channel A to *±5V* and Channel B to *±20V* (if they're not already).

37. Set the scope's *timebase* to *200µs/div* (if it's not already).

38. Use the PicoScope's measurement tools to obtain the peak-to-peak voltage of the signals on the Buffer module's input and output. Record your measurements in Table 3 below.

Table 3

Input voltage	Output voltage

The measure of how much bigger an amplifier's output voltage is compared to its input voltage is called *voltage gain* (A_V). An amplifier's gain can be expressed as a simple ratio and is calculated using the equation:

$$A_V = \frac{Vout}{Vin}$$

Importantly, if the amplifier's output signal is upside-down compared to its input then a negative sign is usually put in front of the gain figure to highlight this fact.

Question 2
Calculate the Buffer module's gain (on its present gain setting).

_____

_____

Ask the instructor to check your work before continuing.

The Buffer module's gain is variable. Usefully, it can be set so that the output voltage is smaller than the input voltage. This is *attenuation*. The next part of the experiment shows how attenuation affects the gain figure.

39. Turn the Buffer module's *Gain* control fully anti-clockwise.

40. Set the vertical scale of Channel B to *±500mV*.

41. Turn the Buffer module's *Gain* control clockwise just a little until you have a small sinewave.

42. Use the PicoScope's measurement tools to obtain the peak-to-peak voltage of the Buffer module's input and output signals. Record your measurements in Table 4 below.

Table 4

Input voltage	Output voltage

Question 3
Calculate the Buffer module's new gain.

_____

_____

Question 4
What's the difference between amplification and attenuation in terms of their gain figure?

_____

_____

_____

Ask the instructor to check your work before continuing.

Amplifiers work by taking the DC power supply voltage and using it to make a copy of the amplifier's input signal. Obviously then, the DC power supply limits the size of the amplifier's output. If the amplifier is forced to try to output a signal that is bigger than the DC power supply voltages, the tops and bottoms of the signal are chopped off. This type of signal distortion is called *clipping*.

Clipping usually occurs when the amplifier's input signal is too big for the amplifier's gain. When this happens, the amplifier is said to be *overdriven*. It can also occur if the amplifier's gain is too big for the input signal. To demonstrate clipping:

43. Set the vertical scale of Channel B to ±20V.

44. Turn the Buffer module's *Gain* control fully clockwise.

Question 5
What do you think the output signal would look like if the buffer's gain was sufficiently large?

Ask the instructor to check your work before continuing.

45. **Turn the Buffer module's *Gain* control fully anti-clockwise.**

Headphones are low impedance devices - typically around 50Ω. Most electronic circuits are not designed to have such low impedances connected to their output. For this reason, headphones should not be directly connected to the output of most of the modules on the Emona Telecoms-Trainer 101.

However, the Buffer module has been specifically designed to handle low impedances. So, it can act as an interface between the modules' outputs and the headphones to let you listen to signals. The next part of the experiment shows how this is done.

46. Ensure that the Buffer module's *Gain* control is turned fully anti-clockwise.

47. Without wearing the headphones, plug them into the Buffer module's headphone socket.

48. Put the headphones on.

49. Turn the Buffer module's *Gain* control clockwise until you can hear the signal and the sound level is comfortable then listen to signal.

50. Disconnect the Buffer module's input from the Master Signals module's *2kHz SINE* output and connect it to the Speech module's output.

51. Speak into the microphone and listen to the signal.

52. Disconnect the Buffer module's input from the Speech module's output and connect it to the Master Signals module's *100kHz SINE* output.

53. Carefully turn the Buffer module's *Gain* control clockwise to see whether you can hear this sinewave but don't turn it fully clockwise.

Question 6
Why is the Master Signals module's *100kHz SINE* output inaudible?

_____

_____

54. Turn the Buffer module's *Gain* control fully anti-clockwise again.

Ask the instructor to check your work before finishing.

2.2 - The Adder and Phase Shifter modules

The Adder module

Several communications and telecommunications systems require that signals be added together. The Adder module has been designed for this purpose.

Procedure

1. If you're starting the experiment from here, gather the equipment listed on page 2-2 (but not the headphones).

2. Locate the Adder module and turn its *g* control (for *Input B*) fully anti-clockwise.

3. Set the Adder module's *G* control for (*Input A*) to about the middle of its travel.

4. Connect the set-up shown in Figure 1 below.

   Note: Although not shown, insert the black plugs of the oscilloscope leads into a ground (*GND*) socket.

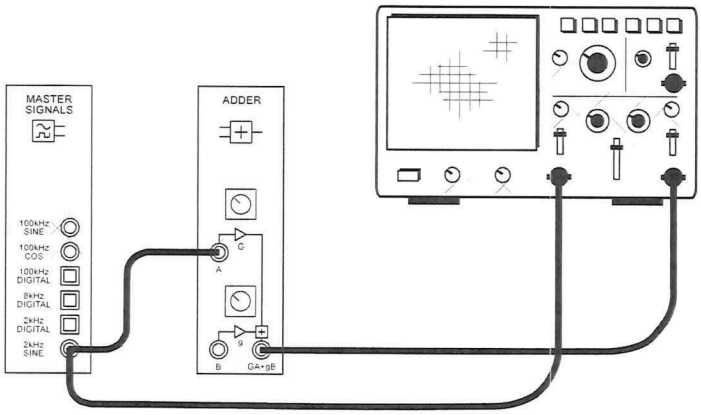


Figure 1

The set-up in Figure 1 on the previous page can be represented by the block diagram in Figure 2 below.

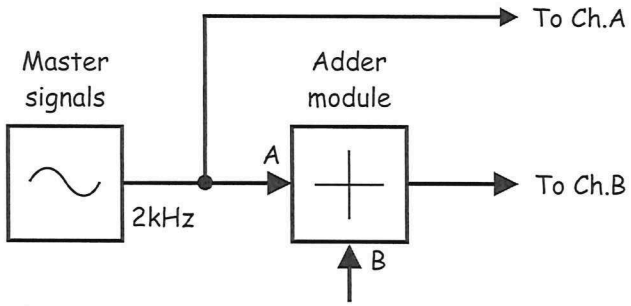


Figure 2

5. Activate the scope's *Auto Setup*.

6. Set the scope's Trigger to *Auto* (if it's not already).

7. Turn on the scope's Channel B (if it's not already).

8. Set the vertical scale of both channels to ±5V.

9. Vary the Adder module's *G* control left and right and observe the effect.

Question 1
What aspect of the Adder module's performance does the *G* control vary?

_____

10. Use the scope's measurements tools to obtain the peak-to-peak voltage on the Adder module's *Input A*. Record the measurement in Table 1 on the next page.

11. Turn the Adder module's *G* control fully clockwise.

12. Use the scope's measurements tool to obtain and record the peak-to-peak voltage on the Adder module's output.

13. Calculate and record the voltage gain of the Adder module's *Input A*.

14. Turn the Adder module's *G* control fully anti-clockwise.

15. Repeat steps 12 and 13.

Table 1		Input voltage	Output voltage	Gain
Input A	Maximum			
	Minimum			

Question 2
What is the range of gains for the A input?

_____

Ask the instructor to check your work before continuing.

16. Turn the Adder module's *G* control fully anti-clockwise.

17. Disconnect the Master Signals module's *2kHz SINE* output from the Adder module's *Input A* and connect it to the Adder's *Input B*.

18. Turn the Adder module's *g* control fully clockwise.

19. Repeat steps 12 and 13 recording your measurement and calculation in Table 2 below.

20. Turn the Adder module's *g* control fully anti-clockwise.

21. Repeat steps 12 and 13.

Table 2		Input voltage	Output voltage	Gain
Input B	Maximum	See Table 1		
	Minimum			

Question 3
Compare the results in Tables 1 and 2. What can you say about the Adder module's two inputs in terms of their gain?

Ask the instructor to check your work before continuing.

22. Turn both of the Adder module's gain controls fully clockwise.

23. Connect the Master Signals module's *2kHz SINE* output to both of the Adder module's inputs.

24. Set the Channel B vertical scale to *±10V*.

Table 3

Adder's output voltage	

Question 4
What is the relationship between the amplitude of the signals on the Adder module's inputs and output?

Ask the instructor to check your work before continuing.

25. Turn the Adder module's *G* control to about the 9 o'clock position.

26. Modify the set-up as shown in Figure 3 below.

Note: The dotted line shows a connection that already exists and need not be changed.

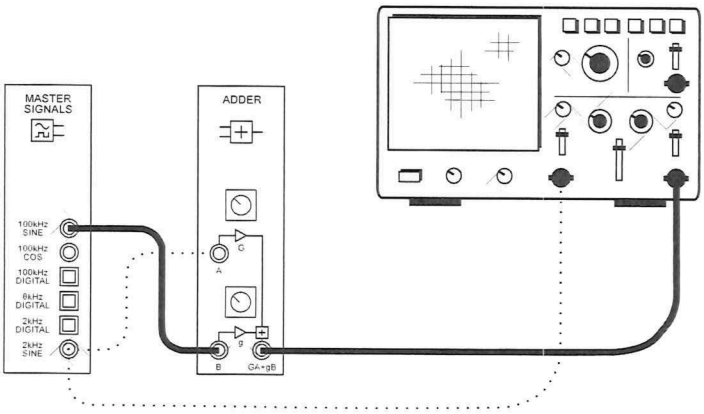


Figure 3

This set-up can be represented by the block diagram in Figure 4 below. The Adder module's output is the simple sum of the two sinewaves at every instant in time.

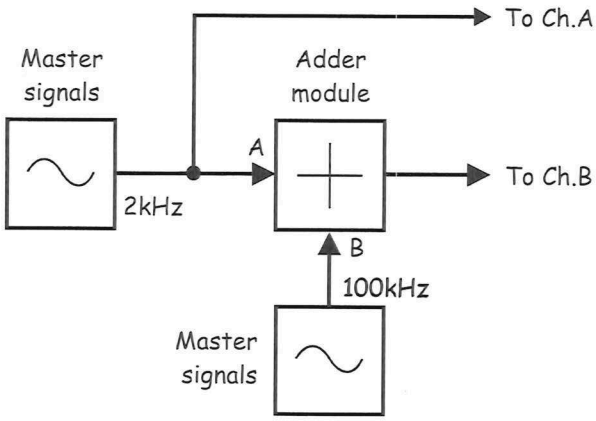


Figure 4

27. Take a moment to think about what you're seeing on the scope's display.

28. Set the scope's timebase to *50µs/div* and take another moment to think about what you're seeing on the scope's display.

Question 5
How many sinewaves do you think there are on the Adder module's output?

_____

_____

29. Activate the Picoscope's Spectrum Mode and set the Spectrum Range to 195kHz.

30. Use a ruler to verify your answer to the previous question.

Ask the instructor to check your work before continuing.

The Phase Shifter module

Several communications and telecommunications systems require that the signal to be transmitted (speech, music and/or video) is phase shifted. Crucial to being able to implement these systems in later experiments is the ability to phase shift any signal by almost any amount up to 360°. The Phase Shifter module has been designed for this purpose.

31. Return to Scope Mode.

32. Set the scope's timebase to *200µs/div.*

33. Locate the Phase Shifter module and set its *Phase Change* control to the *0°* position.

34. Set the Phase Shifter module's *Phase Adjust* control to about the middle of its travel.

35. Connect the set-up shown in Figure 3 below.

    Note 1: Insert the black plugs of the oscilloscope leads into a ground (*GND*) socket.

    Note 2: The LED on the Phase Shifter module will turn on but don't be concerned by this. The LED is used to indicate that the module has automatically adjusted itself for your low frequency input.

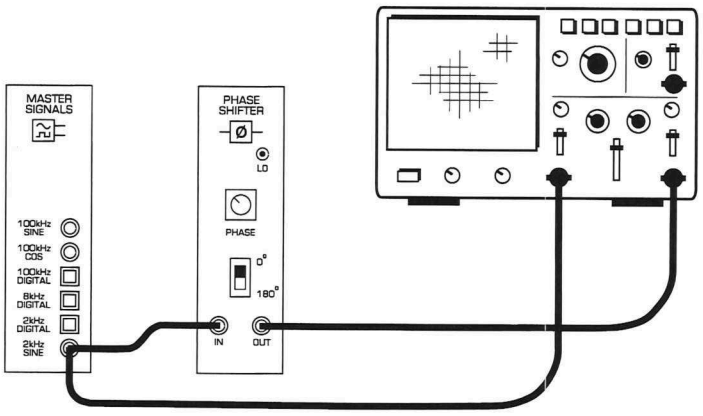


Figure 5

The set-up in Figure 3 on the previous page can be represented by the block diagram in Figure 4 below.

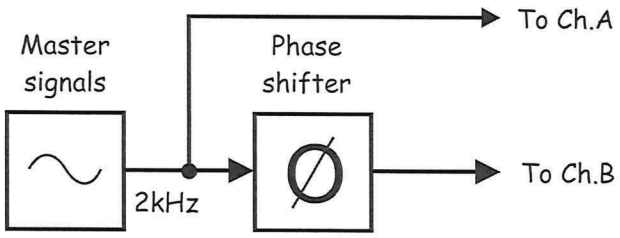


Figure 6

36. Vary the Phase Shifter module's *Phase Adjust* control left and right and observe the effect on the two signals.

37. Set the Phase Shifter module's *Phase Change* control to the *180°* position.

38. Vary the Phase Shifter module's *Phase Adjust* control left and right and observe the effect on the two signals.

Question 6
This module's output signal can be phase shifted by different amounts

- ☐ but it always leads the input signal.
- ☐ but it always lags the input signal.
- ☐ and can either lead or lag the input signal.

Question 7
Is the Phase Shifter module capable of shifting a signal all the way up to 360°? **Tip:** If you're not sure, repeat steps 36 to 38.

Ask the instructor to check your work before finishing.

2.3 - The Voltage Controlled Oscillator (VCO) module

As its name implies, the VCO module is an oscillator like the Master Signals module. However, the VCO module is different in two ways:

- The VCO module has only two outputs (1 analog and 1 digital) whereas the Master Signals module has six (3 analog and 3 digital)
- The VCO module's output frequency is variable whereas the frequency of the Master Signal module's outputs is fixed (at 2kHz, 8kHz and 100kHz)

Procedure

1. If you're starting the experiment from here, gather the equipment listed on page 2-2 (but not the headphones).
2. Locate the VCO module and set its *Gain* control to about the middle of its travel.
3. Set the VCO module's *Frequency Adjust* control to about the middle of its travel.
4. Set the VCO module's *Range* control to the *LO* position.
5. Connect the set-up shown in Figure 1 below.

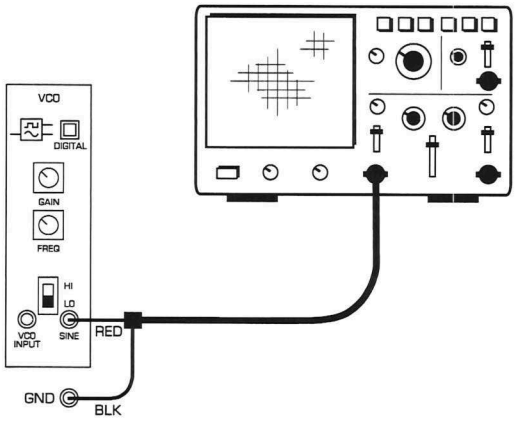


Figure 1

The set-up in Figure 1 on the previous page can be represented by the block diagram in Figure 2 below.

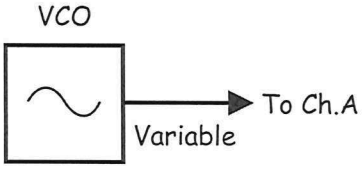


Figure 2

6. Activate the scope's *Auto Setup*.

7. Set the scope's Trigger to *Auto* (if it's not already).

8. Turn off the scope's Channel B for now.

9. Set the vertical scale of Channel A to either *±5V* or *Auto*.

10. Set the scope's *timebase* to *50µs/div*.

11. Use the PicoScope's measurement tools to obtain the peak-to-peak voltage of the signal on the VCO module's *SINE* output. Record the measurement in Table 1 below.

Table 1	Output voltage
SINE output	

Ask the instructor to check your work before continuing.

12. Turn the VCO module's *Frequency Adjust* control fully anti-clockwise but leave the timebase setting on *50µs/div* for the moment.

13. Use the PicoScope's measurement tools to obtain the frequency of the signal on the VCO module's *SINE* output.

    Note: You'll probably find you can't because there isn't at least one full cycle of the waveform being displayed.

14. Readjust the scope's *timebase* to a more appropriate setting for observing the signal.

15. Obtain and record the signal's frequency in Table 1 below.

16. Turn the VCO module's *Frequency Adjust* control fully clockwise and repeat step 15.

17. Set the VCO module's *Range* control to the *HI* position and obtain its minimum and maximum output frequencies.

Table 2	Minimum frequency	Maximum frequency
RANGE on LO		
RANGE on HI		

18. Activate the Picoscope's Spectrum Mode and set the Spectrum Range to *195kHz*.

19. Vary the VCO module's *Frequency Adjust* control and observe what happens to the output on the analyser's display.

Ask the instructor to check your work before continuing.

As you've just seen, the VCO module's output frequency can be changed manually using its *Frequency Adjust* control. The frequency can also be changed electronically by varying the size of a voltage on the module's *VCO INPUT*. The next part of the experiment shows how this is done using the Variable DCV module.

20. Return to Scope Mode.

21. Return the VCO module's *Frequency Adjust* control to about the middle of its travel but leave its *Range* control on *HI*.

22. Locate the Variable DCV module and set its *VDC* control to about the middle of its travel.

23. Connect the set-up shown in Figure 3 below.

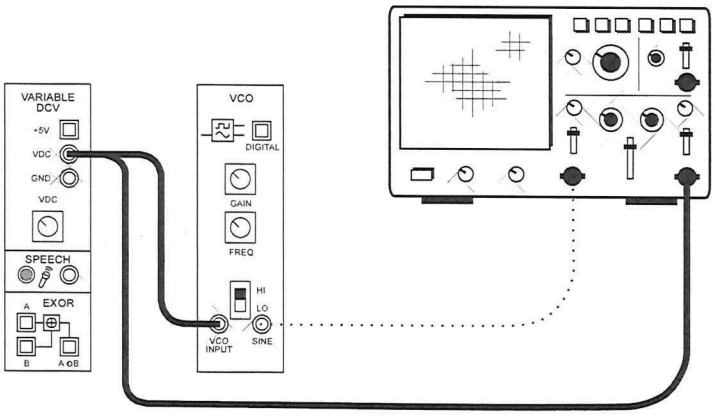


Figure 3

This set-up can be represented by the block diagram in Figure 4 below.

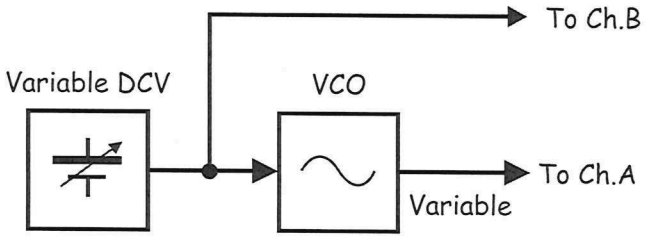


Figure 4

24. Turn on the scope's Channel B by setting its vertical scale to *±5V*.

25. Set the scope's Channel B input coupling to *DC*.

26. Use the scope's measurement tools obtain the *DC Average* voltage on the scope's Channel B input.

27. Turn the Variable DCV module's *VDC* control clockwise while watching the scope's display.

Question 1
What happens to the Variable DCV module's *VDC* output when you turn the VDC control clockwise?

_____

Question 2
What happens to the frequency of the VCO module's output when you turn the VDC control clockwise?

_____

28. Return the Variable DCV module's *VDC* control to about the middle of its travel, then turn it anti-clockwise while watching the scope's display.

Question 3
What happens to the Variable DCV module's *VDC* output when you turn the VDC control anti-clockwise?

_____

Question 4
What happens to the frequency of the VCO module's output when you turn the VDC control anti-clockwise?

_____

Ask the instructor to check your work before finishing.

Name: _____

Class: _____

3 - Using the Telecoms-Trainer 101 to model equations

Experiment 3 - Using the Telecoms-Trainer 101 to model equations

Preliminary discussion
This may surprise you, but mathematics is an important part of electronics and this is especially true for communications and telecommunications. As you'll learn, the output of all communications systems can be described mathematically with an equation.

Although the math that you'll need for this manual is relatively light, there is some. Helpfully, the Emona Telecoms-Trainer 101 can model communications equations to bring them to life.

The experiment
This experiment will introduce you to modelling equations by using the Emona Telecoms-Trainer 101 to implement two relatively simple equations.

It should take you about 45 minutes to complete this experiment.

Equipment

- Emona Telecoms-Trainer 101 (plus power-pack)
- Dual channel 20MHz oscilloscope
- two Emona Telecoms-Trainer 101 oscilloscope leads
- assorted Emona Telecoms-Trainer 101 patch leads

Something you need to know for the experiment

This box contains the definition for an electrical term used in this experiment. Although you've probably seen it before, it's worth taking a minute to read it to check your understanding.

When two signals are **180° out of phase**, they're out of step by half a cycle. This is shown in Figure 1 below. As you can see, the two signals are always travelling in opposite directions. That is, as one goes up, the other goes down (and vice versa).

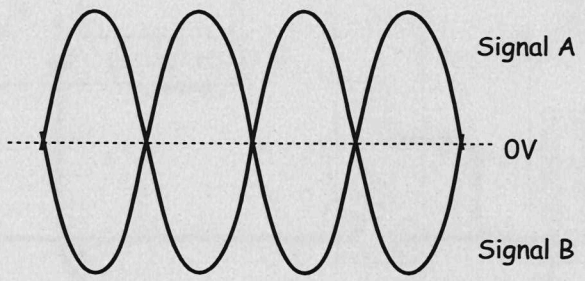


Figure 1

In this part of the experiment, you're going to use the Adder module to add two electrical signals together. Mathematically, you'll be implementing the equation:

Adder module output = Signal A + Signal B

Procedure

1. Gather the equipment listed on page 3-2.

2. Locate the Adder module and set its *G* and *g* controls to about the middle of their travel.

3. Connect the set-up shown in Figure 2 below.

   Note: Although not shown, insert the black plugs of the oscilloscope leads into a ground (*GND*) socket.

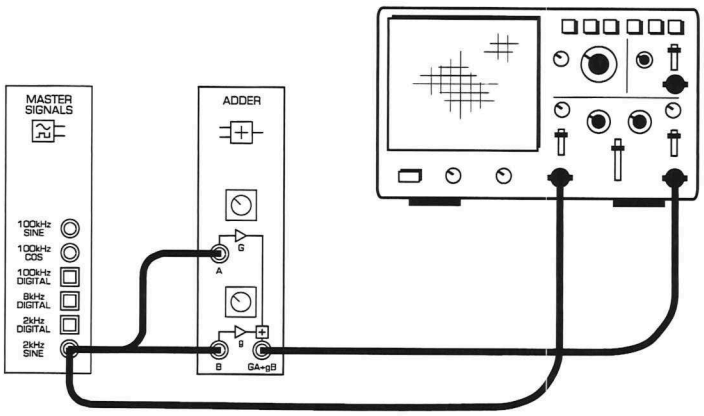


Figure 2

This set-up can be represented by the block diagram in Figure 3 on the next page.

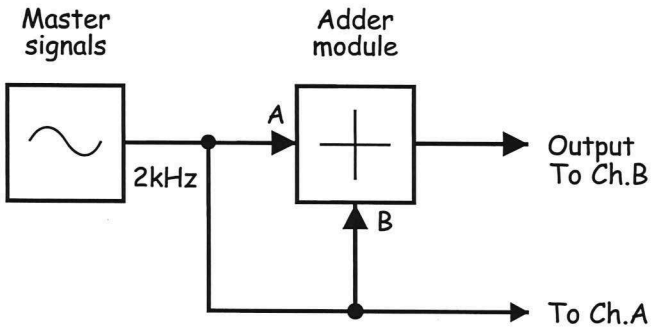


Figure 3

4. Launch the PicoScope app and press *Auto Setup*.

5. Adjust the scope's settings as followings:

   - Vertical scale for Channel A to ±2V
   - Horizontal scale to 200μs
   - Trigger Mode to *Auto*
   - Trigger Channel to *A*

6. Disconnect the lead to the Adder module's *B* input.

7. Use the scope's measurement tools to obtain the peak-to-peak voltage of the Master Signals module's *2kHz SINE* output.

   Record your measurement here: _____

8. Turn on the scope's Channel B and set its vertical scale to ±2V.

9. Adjust the Adder module's *G* control until its output voltage is the same size as its input voltage (measured in Step 7).

   Note 1: Again, use the scope's measurement tools to help with this.

   Note 2: Once done, the gain of the Adder module's A input is -1.

10. Reconnect the lead to the Adder module's *B* input.

11. Disconnect the lead to the Adder module's *A* input.

12. Adjust the Adder module's *g* control until its output voltage is the same size as its input voltage (measured in Step 7).

    Note: Once done, the gain of the Adder module's B input is -1 and means that the Adder module's two inputs should have the same gain.

13. Reconnect the lead to the Adder module's *A* input.

The set-up shown in Figures 3 and 4 is now ready to implement the equation:

Adder module output = Signal A + Signal B

Notice though that the Adder module's two inputs are the same signal - a 2kHz sinewave. Let's assume that the voltage is 4Vp-p (yours will likely be different). With values, the equation above becomes:

Adder module output = 4Vp-p (2kHz sine) + 4Vp-p (2kHz sine)

When the equation is solved, we get:

Adder module output = 8Vp-p (2kHz sine)

Let's see if this is what we get in practice.

14. Remeasure the amplitude of the Master Signals module's *2kHz SINE* output. Record your measurement in Table 1 below.

    Note: The voltage may be a little different to that measured in Step 7 due to loading of the Master Signals module's *2kHz SINE* output.

15. Use this measured value for the two input voltages, together with the equation on the previous page to determine the Adder module's theoretical output voltage.

16. Measure and record the amplitude of the Adder module's output.

Table 1

Measured input voltages	Predicted output voltage	Measured output voltage

Question 1
Is the Adder module's measured output voltage **exactly** as theoretically predicted?

_____

Question 2
What are two reasons for this?

_____

_____

Ask the instructor to check your work before continuing.

In the next part of the experiment, you're going to add two electrical signals together but one of them will be phase shifted. Mathematically, you'll be implementing the equation:

Adder module output = Signal A + Signal B (with phase shift)

17. Locate the Phase Shifter module and set its *Phase Change* control to the *180°* position.

18. Set the Phase Shifter module's *Phase Adjust* control about the middle of its travel.

19. Connect the set-up shown in Figure 4 below.

    Note: Insert the black plugs of the oscilloscope leads into a ground (*GND*) socket.

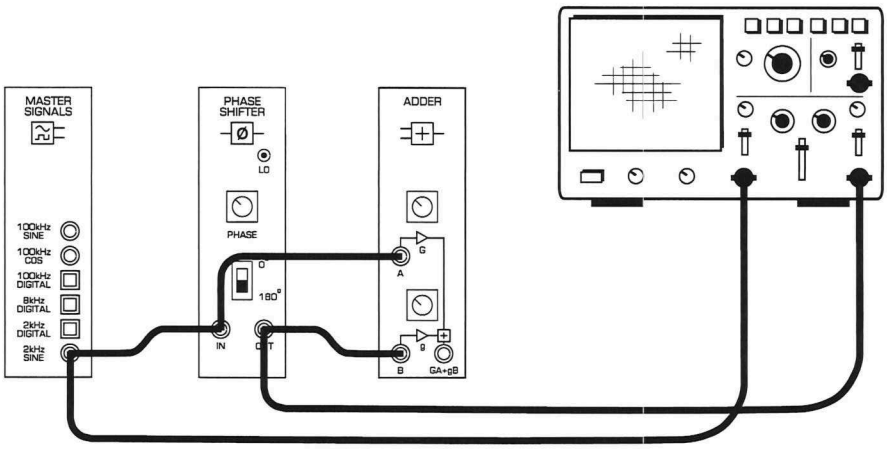


Figure 4

This set-up can be represented by the block diagram in Figure 5 on the next page.

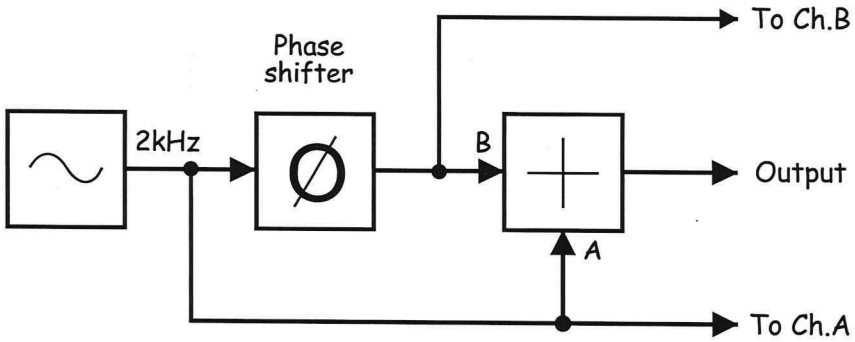


Figure 5

The set-up shown in Figures 4 and 5 is now ready to implement the equation:

Adder module output = Signal A + Signal B (with phase shift)

The Adder module's two inputs are still the same signal – a 2kHz sinewave. So, with values and assuming they're 4Vp-p, the equation is:

Adder module output = 4Vp-p (2kHz sine) + 4Vp-p (2kHz sine with phase shift)

As the two signals have the same amplitude and frequency, if the phase shift is exactly 180° then their voltages at any point in the waveform is always exactly opposite. That is, when one sinewave is +1V, the other is -1V. When one is +3.75V, the other is -3.75V and so on. This means that, when the equation above is solved, we get:

Adder module output = 0Vp-p

Let's see if this is what we get in practice.

20. Adjust the Phase Shifter module's *Phase Adjust* control until the two signals look like they're 180° out of phase with each other.

21. Disconnect the scope's Channel B lead from the Phase Shifter module's output and connect it to the Adder module's output.

22. Measure the amplitude of the Adder module's output. Record your measurement in Table 2 below.

    Tip: You'll likely need to set the vertical scale of Channel B to a lower ± value - Just keep setting it to smaller values until you see a signal and ignore any noise that becomes obvious.

Table 2

Output voltage

Question 3
What are two reasons for the output not being 0V as theoretically predicted?

_____

_____

Ask the instructor to check your work before continuing.

The following procedure can be used to adjust the Adder and Phase Shifter modules so that the set-up has a *null* output (that is, where they cancel each other out almost completely and the output is as is close to 0Vp-p as possible).

23. Vary the Phase Shifter module's *Phase Adjust* control left and right and observe the effect on the Adder module's output.

24. Adjust the Phase Shifter module's *Phase Adjust* control to obtain the smallest output voltage.

Question 5
What can be said about the phase shift between the signals on the Adder module's two inputs now?

_____

25. Vary the Adder module's *g* control left and right and observe the effect on the Adder module's output.

26. Adjust the Adder module's *g* control to obtain the smallest output voltage.

Question 6
What can be said about the gain of the Adder module's two inputs now?

_____

Ask the instructor to check your work before finishing.

Name: _____

Class: _____

4 - Amplitude modulation (AM)

Experiment 4 - Amplitude modulation (AM)

Preliminary discussion

In an amplitude modulation (AM) communications system, speech and music are converted into an electrical signal using a device such as a microphone. This electrical signal is called the *message* or *baseband* signal. The message signal is then used to electrically vary the amplitude of a pure sinewave called the *carrier*. The carrier usually has a frequency that is much higher than the message's frequency.

Figure 1 below shows a simple message signal and an unmodulated carrier. It also shows the result of amplitude modulating the carrier with the message. Notice that the modulated carrier's amplitude varies above and below its unmodulated amplitude.

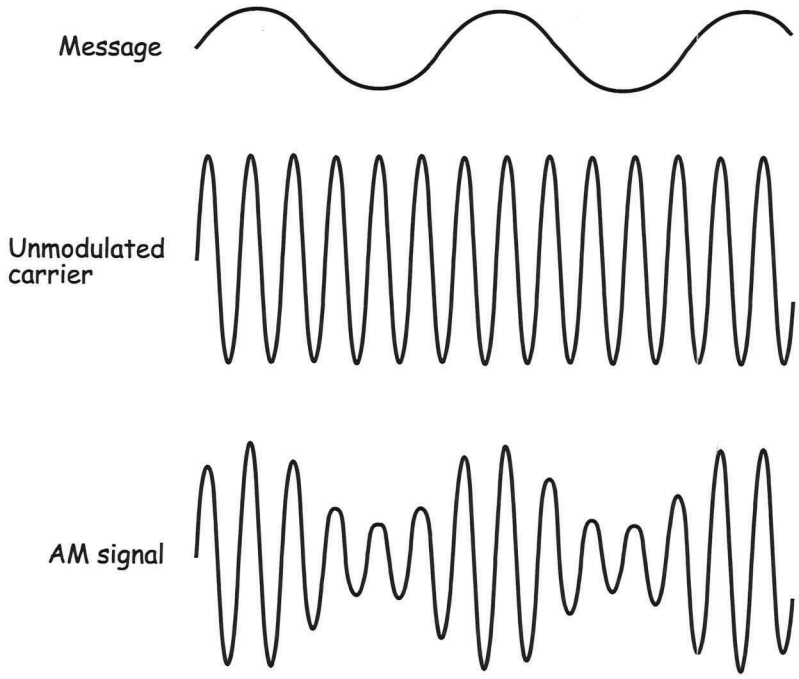


Figure 1

© 2008 Emona Instruments

Figure 2 below shows the AM signal at the bottom of Figure 1 but with a dotted line added to track the modulated carrier's positive peaks and negative peaks. These dotted lines are known in the industry as the signal's *envelopes*. If you look at the envelopes closely you'll notice that the upper envelope is the same shape as the message. The lower envelope is also the same shape but upside-down (inverted).

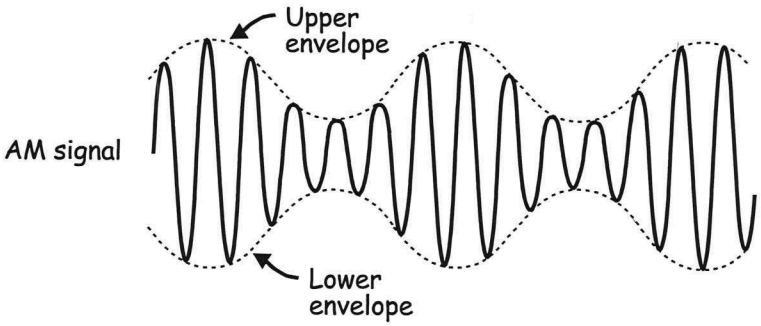


Figure 2

In telecommunications theory, the mathematical model that defines the AM signal is:

AM = (DC + message) × the carrier

When the message is a simple sinewave (like in Figure 1) the equation's solution (which necessarily involves some trigonometry that is not shown here) tells us that the AM signal consists of three sinewaves:

- One at the carrier frequency
- One with a frequency equal to the sum of the carrier and message frequencies
- One with a frequency equal to the difference between the carrier and message frequencies

In other words, for every sinewave in the message, the AM signal includes a pair of sinewaves - one above and one below the carrier's frequency. Complex message signals such as speech and music are made up of thousands sinewaves and so the AM signal includes thousands of pairs of sinewaves straddling carrier. These two groups of sinewaves are called the *sidebands* and so AM is known as *double-sideband, full carrier* (DSBFC).

Importantly, it's clear from this discussion that the AM signal doesn't consist of any signals at the message frequency. This is despite the fact that the AM signal's envelopes are the same shape as the message.

 © 2008 Emona Instruments

The experiment

In this experiment you'll use Emona Telecoms-Trainer 101 to generate a real AM signal by implementing its mathematical model. This means that you'll add a DC component to a pure sinewave to create a message signal then multiply it with another sinewave at a higher frequency (the carrier). You'll examine the AM signal using the scope and compare it to the original message. You'll do the same with speech for the message instead of a simple sinewave.

Following this, you'll vary the message signal's amplitude and observe how it affects the modulated carrier. You'll also observe the effects of modulating the carrier too much. Finally, you'll measure the AM signal's depth of modulation using a scope.

It should take you about 1 hour to complete this experiment.

Equipment

- Emona Telecoms-Trainer 101 (plus power-pack)
- Dual channel 20MHz oscilloscope
- two Emona Telecoms-Trainer 101 oscilloscope leads
- assorted Emona Telecoms-Trainer 101 patch leads

Procedure

Part A - Generating an AM signal using a simple message

1. Gather a set of the equipment listed above.

2. Locate the Adder module and turn its *G* and *g* controls fully anti-clockwise.

3. Locate the Variable DCV module and turn its *DC Voltage* control almost fully anti-clockwise.

 © 2008 Emona Instruments

4. Connect the set-up shown in Figure 3 below using the PicoScope for the oscilloscope.

Note: Insert the black plug of the oscilloscope lead into a ground (*GND*) socket.

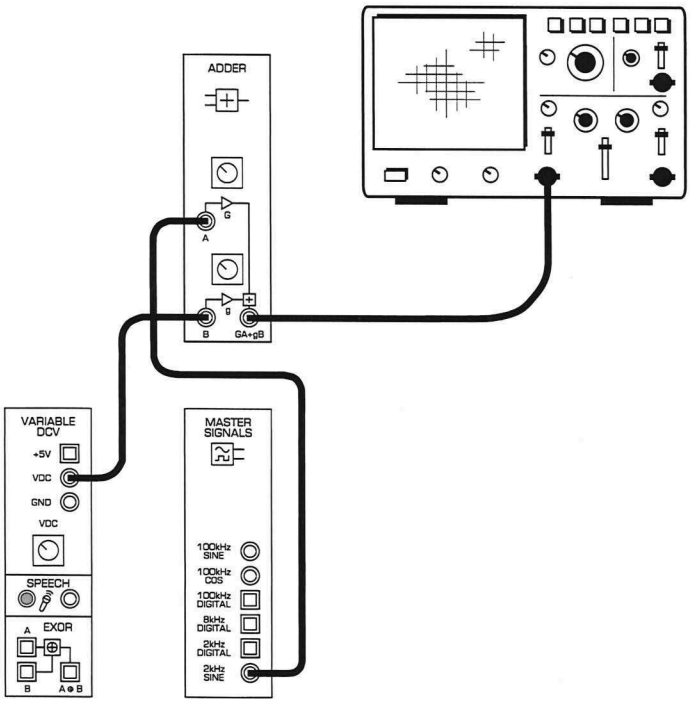


Figure 3

This set-up can be represented by the block diagram in Figure 4 below. It implements the highlighted part of the equation: *AM* = **(DC + message)** × the carrier.

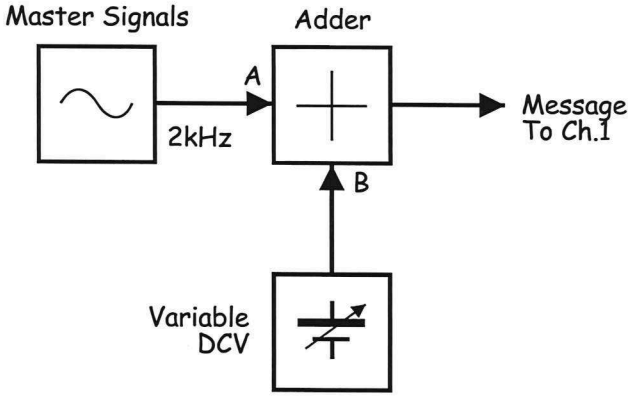


Figure 4

 © 2008 Emona Instruments

At the moment, the scope should just be showing a flat trace because the Adder module's output is 0V. Once you've completed steps 5 to 8 below, the AM equation will have the following values:

AM = (1VDC + 1Vp-p 2kHz sine) × the carrier

5. Launch the PicoScope app and press *Auto Setup*.

6. Adjust the scope's settings as followings:

   - Vertical scale for Channel *A* to *±2V*
   - Horizontal scale to *200µs*
   - Trigger Mode to *Auto*
   - Trigger Channel to *A*
   - Trigger Threshold to *1V*

7. While watching the Adder module's output on the scope, turn its *g* control clockwise until the DC level is 1V.

   Note: You can do this by visual inspection to get it close but use the scope's measurement tools for the most accurate output voltage.

8. While watching the Adder module's output on the scope, turn its *G* control clockwise to obtain a 1Vp-p sinewave.

   Note: Again, use the scope's measurement tools for the most accurate output voltage.

Ask the instructor to check your work before continuing.

 © 2008 Emona Instruments

9. Modify the set-up as shown in Figure 5 below.

> **Before you do...**
> The set-up in Figure 5 below builds on the set-up that you've already wired so don't pull it apart. To highlight the changes that we want you to make, we've shown your existing wiring as dotted lines.

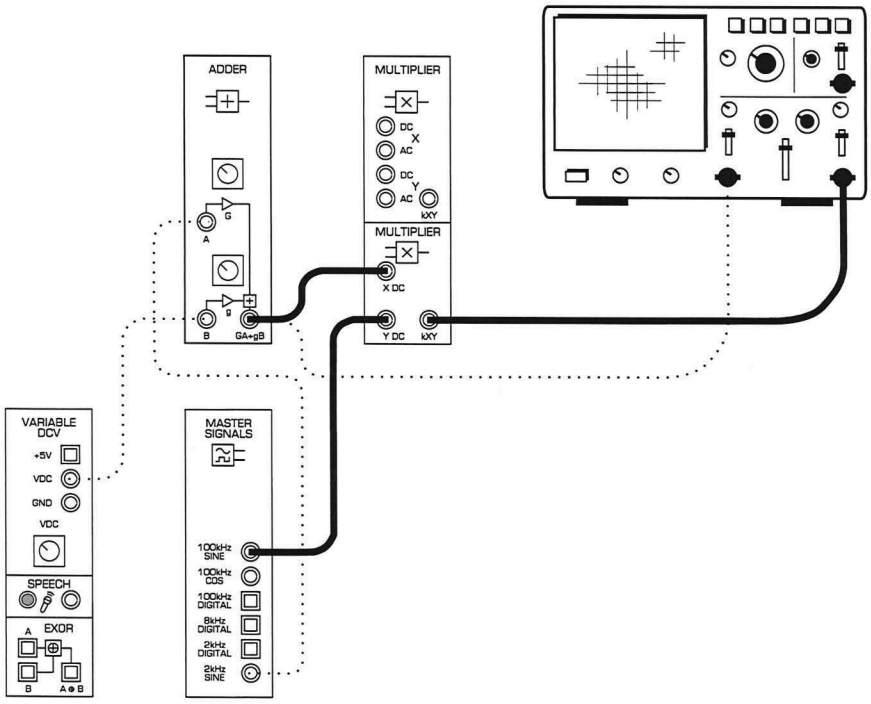


Figure 5

10. Turn on the scope's Channel B and set its vertical scale to ±5V.

    Note 1: Once done, you should see an AM signal on the scope's display much like the example in Section 4 of the workbook.

    Note 2: Be sure to read the notes on the next page before calling the instructor to check your work.

 © 2008 Emona Instruments

The set-up in Figure 5 can be represented by the block diagram in Figure 6 below. The additions that you've made to the original set-up implement the highlighted part of the equation:
AM = (DC + message) × **the carrier**.

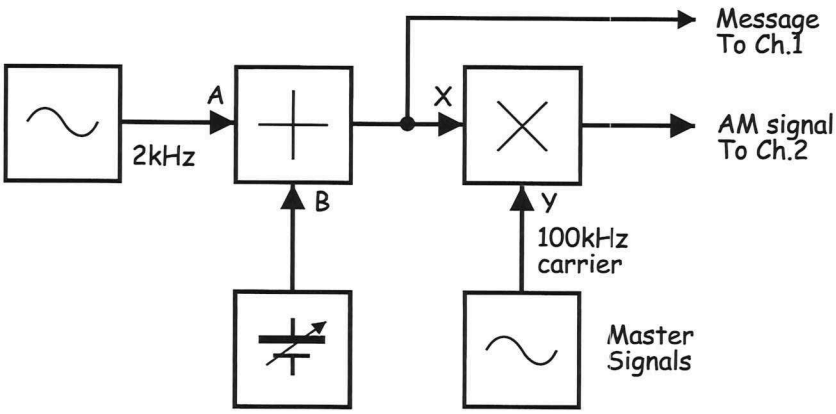


Figure 6

With values, the equation is:

AM = (1VDC + 1Vp-p 2kHz sine) × 4Vp-p 100kHz sine

Ask the instructor to check your work before continuing.

11. Use the scope's Channel A *Vertical Position* control to overlay the message with the AM signal's envelopes and compare them.

Question 1
What feature of the Multiplier module's output suggests that it's an AM signal? **Tip:** If you're not sure about the answer to the questions, see the preliminary discussion.

 © 2008 Emona Instruments

Question 2
The AM signal is a complex waveform consisting of more than one sinewaves. Is one of the sinewaves at the message frequency (ie at 2kHz)? Explain your answer.

_____

_____

Question 3
For the given inputs to the Multiplier module, how many sinewaves does the AM signal consist of, and what are their frequencies?

_____

_____

12. Set the PicoScope to Spectrum Mode.

13. Turn off Channel A and turn on Channel B.

14. Set the Spectrum Range to 200kHz.

    Note: Once done, the spectrum will appear to have many significant sinewaves. However, this is misleading because the vertical scale is logarithmic. Most of them are actually quite small compared to the sinewaves in the AM signal. (To explain, a sinewave that is -20dB relative to another is one-tenth its size. A sinewave that is -40dB relative to another is one-hundredth its size.)

15. Examine the AM signal's spectral composition.

    Note 1: You may find it helpful to use the PicoScope's vertical cursors to identify the exact frequency of the AM signal's significant spectral components.

    Note 2: Once done, make any necessary changes to your answers to Questions 2 and 3.

Ask the instructor to check your work before continuing.

 © 2008 Emona Instruments

Part B - Generating an AM signal using speech

This experiment has generated an AM signal using a sinewave as a model for the message. However, the message in commercial AM communications systems is much more likely to be speech and music. The next part of the experiment lets you see what an AM signal looks like when modulated by speech.

16. Return the PicoScope to Scope Mode.

17. Set the scope's horizontal scale to *2ms*.

18. Disconnect the plug on the Master Signals module's *2kHz SINE* output that connects to the Adder module's *A* input.

19. Connect it to the Speech module's output as shown in Figure 7 below.

    Remember: Dotted lines show leads already in place.

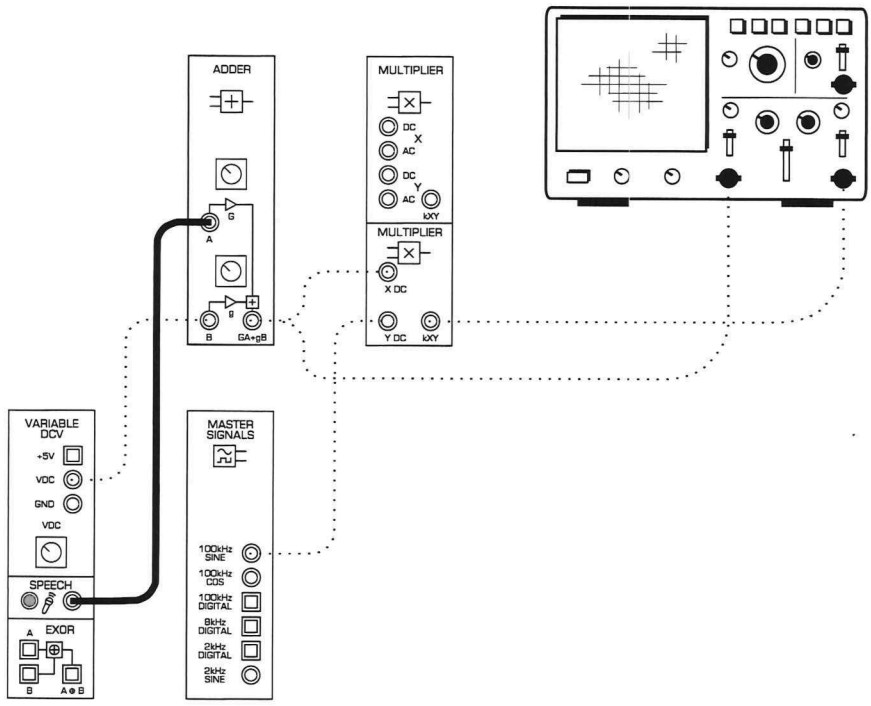


Figure 7

20. Talk, hum and whistle while watching the scope's display.

    Note: The gain of the Speech module's amplifier is quite high so you'll find that it's easy to overdrive it, causing clipping on the message which is observed on the AM signal too. Try to not talk/hum/etc too loudly to avoid this problem.

 © 2008 Emona Instruments

21. Locate the scope's stop button and hover the mouse's cursor over it (but don't click on it yet).

22. Talk or hum again, but this time, stop the scope's acquisition (to freeze the display) when your speech produces a wide variety of voltages in the message signal.

    Tip: Try varying the volume of your voice as you talk to achieve this effect (without overdriving the Speech module's amp).

23. Overlay the message with the AM signal's envelopes again to compare them.

    Note: You can probably see now why a sinewave is used for a model message when teaching and writing about AM signals!

Ask the instructor to check your work before continuing.

24. Restart the scope (if you haven't already).

25. Be as quiet as you can why watching the message and AM signals.

    Question 4
    Why are the AM signal's envelopes still not flat?

    _____

    _____

26. Disconnect the plug on the Speech module's output.

27. Connect the plug to one of the trainer's ground (GND) sockets.

    Note: The AM signal's envelopes should now be flat.

 © 2008 Emona Instruments

Question 5
Why does the AM signal still have a sinewave when there is no message signal? **Tip:** If you're not sure, look at the signal in Spectrum Mode.

Ask the instructor to check your work before continuing.

Part C – Investigating depth of modulation
It's possible to modulate the carrier by different amounts. This part of the experiment let's you investigate this.

28. If you set the PicoScope to Spectrum Mode to help answer Question 5, return it to Scope Mode.

29. Return the scope's horizontal scale to *200µs*.

30. Disconnect the plug that you just connected to ground and connect it to the Master Signals module's *2kHz SINE* output again.

31. Vary the message signal's amplitude a little by turning Adder module's *G* control left and right and notice the effect on the AM signal.

Question 6
What is the relationship between the message's amplitude and the amount of the carrier's modulation?

 © 2008 Emona Instruments

Ask the instructor to check your work before continuing.

The extent that a message modulates a carrier is known as the *modulation index* (m). Modulation index is an important characteristic of an AM signal for several reasons including calculating the distribution of the signal's power between the carrier and sidebands.

Figure 8 below shows two key dimensions of an amplitude modulated carrier. These two dimensions allow a carrier's modulation index to be determined without knowing the size of the message or carrier signals.

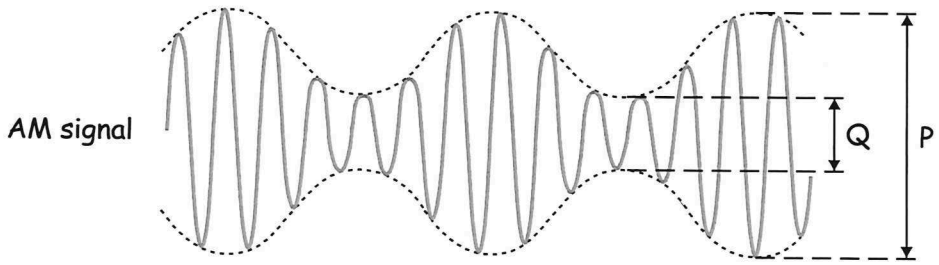


Figure 8

The next part of the experiment lets you practise measuring these dimensions to calculate a carrier's modulation index.

32. Adjust the Adder module's *G* control to return the message signal's amplitude to 1Vp-p.

33. Use the PicoScope's cursors to measure the AM signal's *B* dimension. Record your measurement in Table 1 below.

34. Measure and record the AM signal's *A* dimension.

35. Calculate and record the AM signal's depth of modulation using the equation for this in the workbook.

Table 1

B dimension	A dimension	m

 © 2008 Emona Instruments

Ask the instructor to check your work before continuing.

A problem that is important to avoid in AM transmission is *over-modulation*. When the carrier is over-modulated, it can upset the receiver's operation. The next part of the experiment gives you a chance to observe the effect of over-modulation.

36. Set the vertical scale of Channel A to *±5V* and the vertical scale of Channel B to *±10V*.

37. Increase the message signal's amplitude by turning the Adder module's *G* control to about the 12 o'clock postion.

    Note: This should cause the AM signal's carrier to be over-modulated.

38. Use the scope's Channel A *Vertical Position* control to overlay the message with the AM signal's envelopes and compare them.

Question 7
Over-modulation can cause two problems. One is "sideband splatter" where the AM signal produces spectral components that can interfere with adjacent channels. What's the other problem?

_____

_____

_____

Question 8
What do you think is a carrier's maximum modulation index without over-modulation?

- ☐ A minus number
- ☐ 0
- ☐ 1
- ☐ Greater than 1

 © 2008 Emona Instruments

Ask the instructor to check your work before finishing.

 © 2008 Emona Instruments

Name: _____

Class: _____

6 - AM demodulation

Experiment 6 - AM demodulation

Preliminary discussion

If you've completed Experiment 4 then you've seen what happens when you use a 2kHz sinewave to amplitude modulate a carrier to produce an AM signal. Importantly, you would have seen a key characteristic of an AM signal - its envelopes are the same shape as the message (though the lower envelope is inverted).

Recovering the original message from a modulated carrier is called *demodulation* and this is the main purpose of communications and telecommunications receivers. The circuit that is widely used to demodulate AM signals is called an *envelope detector*. The block diagram of an envelope detector is shown in Figure 1 below.

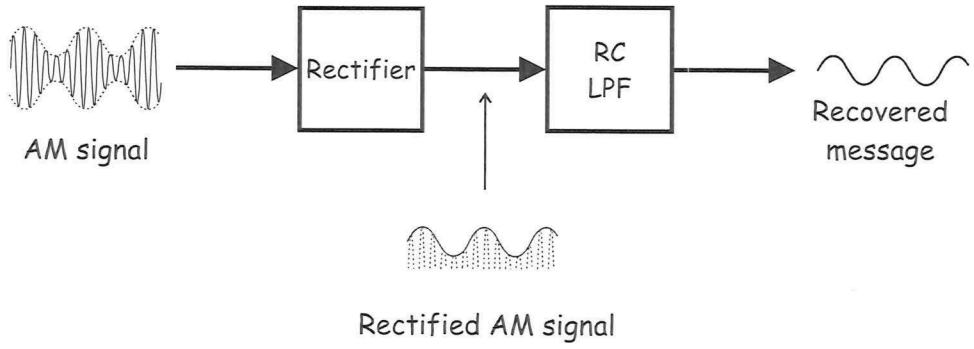


Figure 1

As you can see, the rectifier stage chops the AM signal in half letting only one of its envelopes through (the upper envelope in this case but the lower envelope is just as good). This signal is fed to an RC LPF which tracks the peaks of its input. When the input to the RC LPF is a rectified AM signal, it tracks the signal's envelope. Importantly, as the envelope is the same shape as the message, the RC LPF's output voltage is also the same shape as the message and so the AM signal is demodulated.

A limitation of envelope detector shown in Figure 1 is that it cannot accurately recover the message from over-modulated AM signals. To explain, recall that when an AM carrier is over-modulated the signal's envelope is no-longer the same shape as the original message. Instead, the envelope is distorted and so, by definition, this means that the envelope detector must produce a distorted version of the message.

 © 2008 Emona Instruments

The experiment

In this experiment you'll use the Emona Telecoms-Trainer 101 to generate an AM signal by implementing its mathematical model. Then you'll set-up an envelope detector using the *Rectifier* and *RC LPF* on the trainer's Utilities module.

Once done, you'll connect the AM signal to the envelope detector's input and compare the demodulated output to the original message and the AM signal's envelope. You'll also observe the effect that an over-modulated AM signal has on the envelope detector's output.

It should take you about 1 hour to complete Parts A to D of this experiment.

Equipment

- Emona Telecoms-Trainer 101 (plus power-pack)
- Dual channel 20MHz oscilloscope
- two Emona Telecoms-Trainer 101 oscilloscope leads
- assorted Emona Telecoms-Trainer 101 patch leads
- one set of headphones (stereo)

 © 2008 Emona Instruments

Procedure

Part A – Setting up the AM modulator

To experiment with AM demodulation you need an AM signal. The first part of the experiment gets you to set one up.

1. Gather a set of the equipment listed above.

2. Locate the Adder module and turn its *G* and *g* controls fully anti-clockwise.

3. Locate the Variable DCV module and turn its *DC Voltage* control almost fully anti-clockwise.

4. Connect the set-up shown in Figure 2 below using the PicoScope.

   Note: Insert the black plugs of the oscilloscope leads into a ground (*GND*) socket.

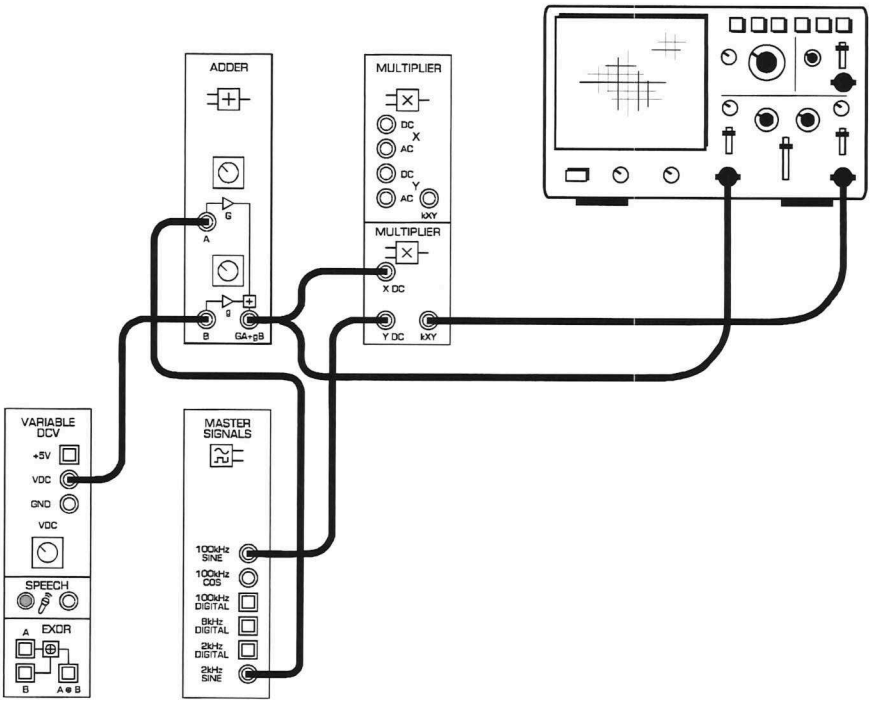


Figure 2

 © 2008 Emona Instruments

The set-up in Figure 2 can be represented by the block diagram in Figure 3 below. It generates a 100kHz carrier that is amplitude modulated by a 2kHz sinewave message.

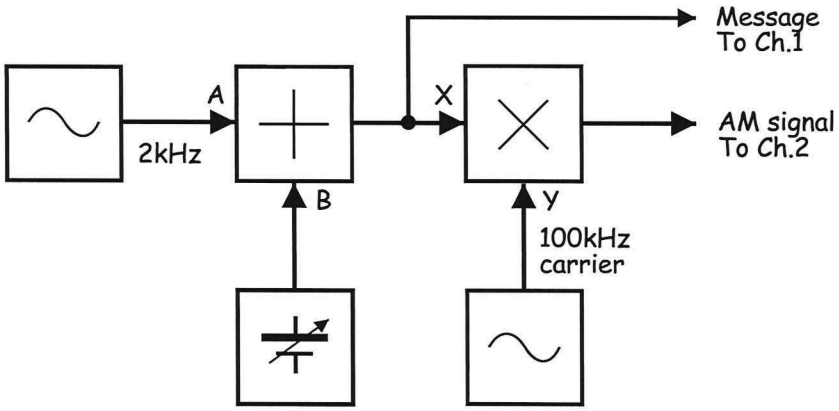


Figure 3

Note: Although the scope is now displaying the Adder module's output, it'll just be a straight line because its output is 0V.

5. Launch the PicoScope app and press *Auto Setup*.

6. Adjust the scope's settings as followings:

   - Vertical scale for Channel A to *±2V*
   - Horizontal scale to *200µs*
   - Trigger Mode to *Auto*
   - Trigger Channel to *A*
   - Trigger Threshold to *1V*

   Note: Leave Channel B turned off for the moment.

7. While watching the Adder module's output on the scope, turn its *g* control clockwise until the DC level is 1V.

   Note: You can do this by visual inspection to get it close enough (there's no need to use the scope's measurement cursors unless you would prefer to).

8. While watching the Adder module's output on the scope, turn its *G* control clockwise to obtain a 1Vp-p sinewave.

   Note: Again, you can do this by visual inspection to get it close enough.

 © 2008 Emona Instruments

9. Turn on the scope's Channel B and set its vertical scale to ±5V.

10. Adjust the vertical positions of the two channels so that they're not overlayed.

11. Verify that the signal on Channel B is an AM signal.

Question 1
How do you know for certain that the signal on Channel B is DSBFC AM?

_____

_____

_____

12. Call the teacher who will adjust the AM signal's depth of modulation to a random value.

13. Measure the AM signal's depth of modulation recording your measurements and calculation in Table 1 below.

Tip: If you're not sure how to do this, see Section 4 of the workbook.

Table 1

B dimension	A dimension	m

Ask the instructor to check your work before continuing.

© 2008 Emona Instruments

Part B – Recovering the message using an envelope detector

14. Modify the set-up as shown in Figure 4 below.

Remember: Dotted lines show leads already in place.

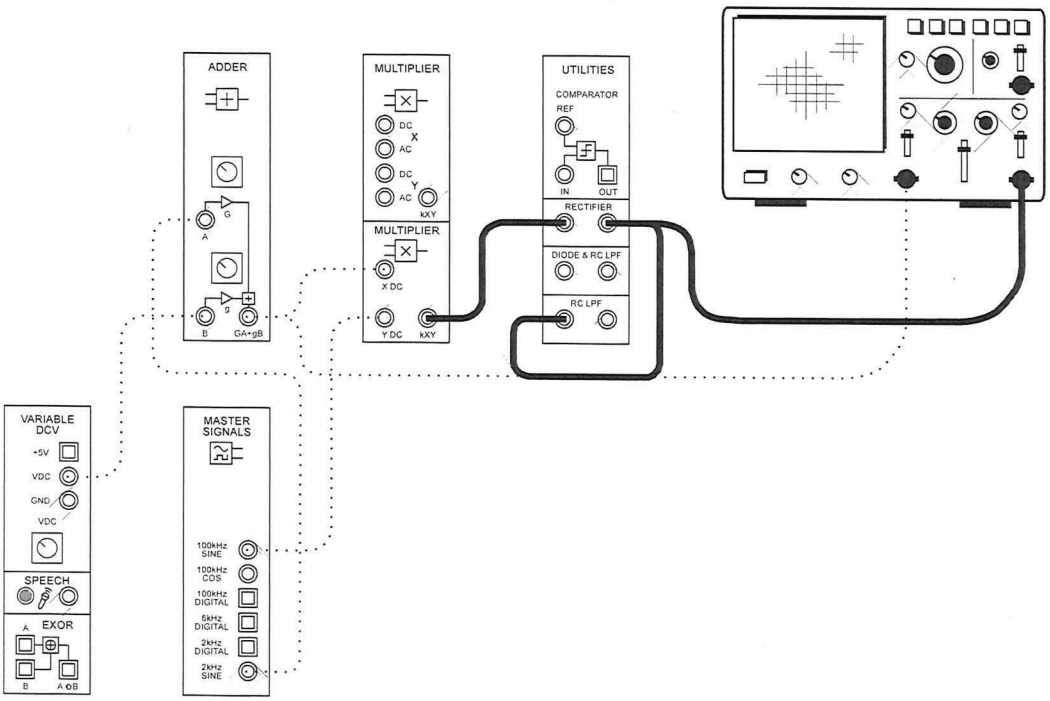


Figure 4

The additions to the set-up in Figure 4 can be represented by the block diagram in Figure 5 below. As you can see, it's the envelope detector explained in the preliminary discussion.

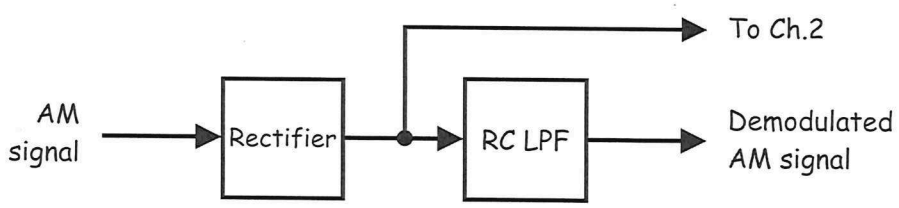


Figure 5

 © 2008 Emona Instruments

Question 2
Explain why the signal out of the rectifier looks the way it does.

15. Disconnect the scope's *Channel B* input from the *Rectifier's* output and connect it to the *RC LPF's* output instead.

16. Set the scope's Channel B vertical scale to *±1V*.

Question 3
What is the name for the signal on the RC LPF's output?

Question 4
Briefly explain why the signal on the RC LPF's output looks like the original message.

Question 5
The original message and the recovered message are out of phase. If you could compare the recovered message out of a receiver with the original message at the transmitter some distance away in a real communications system, do you think these two signals would be in phase or likely out of phase. Give a reason for your answer.

 © 2008 Emona Instruments

Ask the instructor to check your work before continuing.

Part C – Investigating the message's amplitude on the recovered message

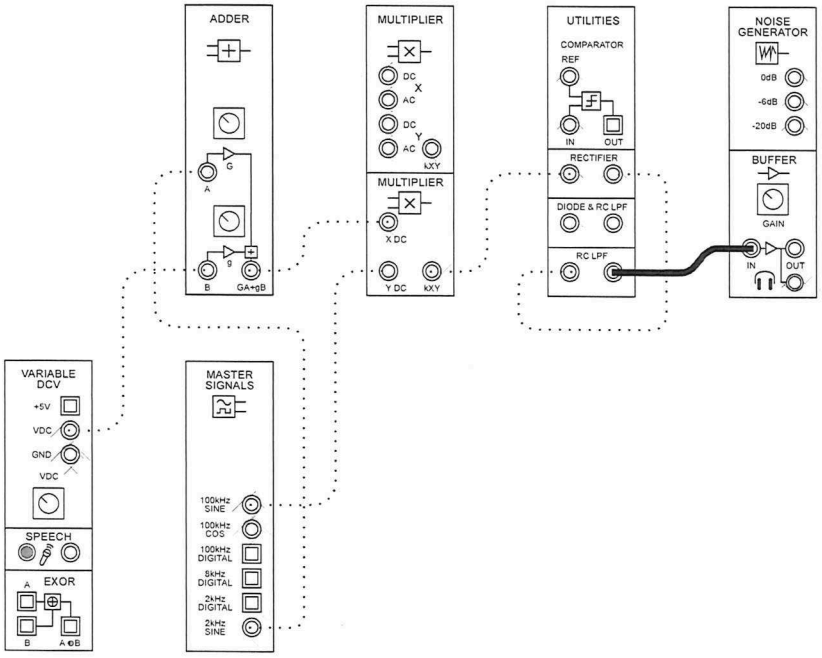


Figure 6

17. Turn the Buffer module's *Gain* control fully anti-clockwise.

18. Without wearing the headphones, plug them into the Buffer module's headphone socket.

19. Put the headphones on.

20. Set the Buffer module's *Gain* control to a comfortable sound level.

21. Vary the message signal's amplitude up and down a little (by turning the Adder module's *G* control left and right a little) while watching and listening to the demodulated signal.

 © 2008 Emona Instruments

Question 6
What is the relationship between the amplitude of the two message signals?

_____

_____

22. Slowly increase the message signal's amplitude by turning the Adder module's *G* control to about the 12 o'clock position.

    Note: As you do, you'll likely need to change the Channel A and B vertical scales to ±5V.

23. Look closely at, and listen to, the demodulated signal (ie the recovered message).

    Note: You should notice a visual and audible distortion of the recovered message.

Question 7
What do you think is causing the heavy distortion of the recovered message?

_____

24. To verify your answer to Question 7 connect the scope's Channel A input to the AM modulator's output and set its *Trigger Source* control to *CHB*.

Question 8
Why does over-modulation cause this distortion in the recovered message?

_____

_____

_____

Ask the instructor to check your work before continuing.

 © 2008 Emona Instruments

Part D – Transmitting and recovering speech/music using AM

This experiment has set up an AM communication system to "transmit" a message that is a 2kHz sinewave. The next part of the experiment lets you use the set-up to modulate, transmit, demodulate and listen to speech.

25. Reconnect the scope's Channel A to the Adder module's output, its *Trigger Source* to *CHA* and the vertical scales for Channel A and B to *±2V* and *±1V* respectively.

26. Adjust the message signal's amplitude back to 1Vp-p (by turning the Adder module's *G* control anti-clockwise).

27. Modify the set-up as shown in Figure 7 below.

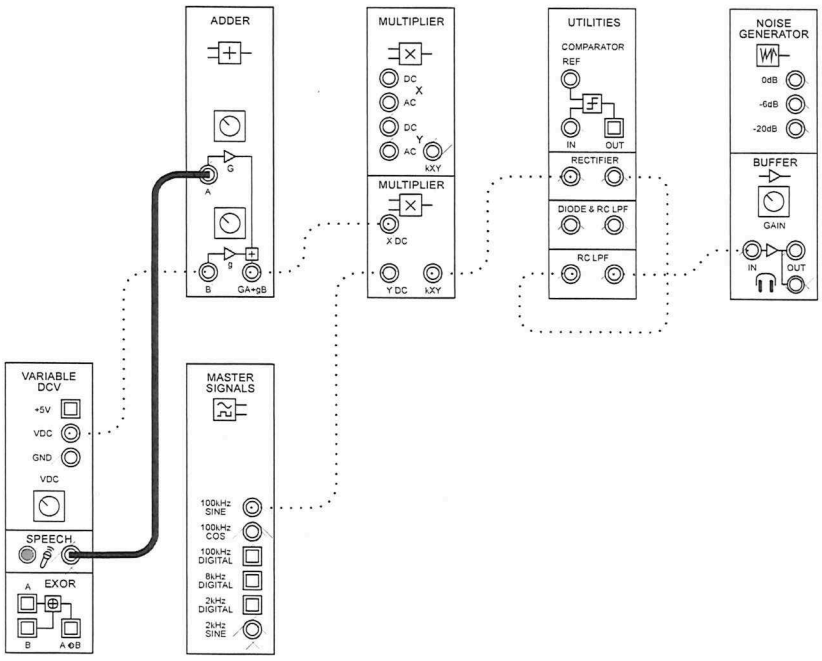


Figure 7

28. Set the scope's *Timebase* control to the *2ms/div* position.

29. Talk, hum and whistle while watching the scope's display and listening on the headphones.

 © 2008 Emona Instruments

A problem with using the sounds that you make to test this arrangement is the fact that you can hear yourself talk, etc through headphones and even though your own body. That makes it difficult to separate what you can hear this way from what you can hear through this AM communication system. So, to mitigate that problem...

30. Set up your phone to output music very quietly to its speaker and place the speaker next to the trainer's microphone. Listen to this music via the headphones.

Question 9
Why do you think the sound of this music is not as good as out of your phone's speaker?

_____

_____

_____

Ask the instructor to check your work before finishing.

 © 2008 Emona Instruments

Name: _____

Class: _____

9 - Frequency modulation

Experiment 9 - Frequency modulation

Preliminary discussion

A disadvantage of the AM, DSBSC and SSB communication systems is that they are susceptible to picking up electrical noise in the transmission medium (the *channel*). This is because noise changes the amplitude of the transmitted signal and the demodulators of these systems are looking for amplitude variations.

As its name implies, frequency modulation (FM) uses a message's amplitude to vary the frequency of a carrier instead of its amplitude. This means that the FM demodulator is designed to look for changes in frequency instead. As such, it is less affected by amplitude variations and so FM is less susceptible to noise. This makes FM a better communications system in this regard.

There are several methods of generating FM signals but they all basically involve an oscillator with an electrically adjustable frequency. The oscillator uses an input voltage to affect the frequency of its output. Typically, when the input is 0V, the oscillator outputs a signal at its *rest* frequency (also commonly called the *free-running* or *centre* frequency). If the applied voltage varies above or below 0V, the oscillator's output frequency deviates above and below the rest frequency. Moreover, the amount of deviation is affected by the amplitude of the input voltage. That is, the bigger the input voltage, the greater the deviation.

Figure 1 below shows a simple message signal (a bipolar squarewave) and an unmodulated carrier. It also shows the result of frequency modulating the carrier with the message.

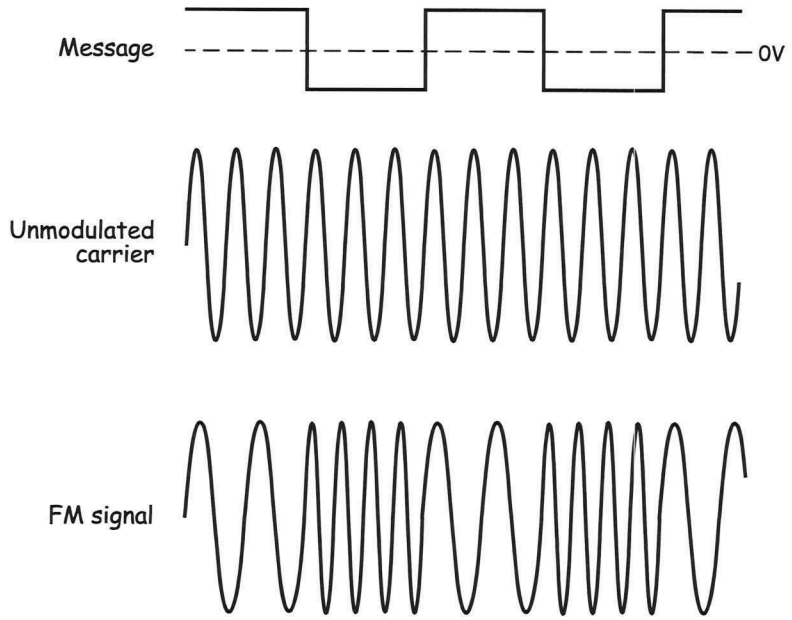


Figure 1

 © 2008 Emona Instruments

There are a few things to notice about the FM signal. First, its envelopes are flat – recall that FM doesn't vary the carrier's amplitude. Second, its period (and hence its frequency) changes when the amplitude of the message changes. Third, as the message alternates above and below 0V, the signal's frequency goes above and below the carrier's frequency. (Note: It's equally possible to design an FM modulator to cause the frequency to change in the opposite direction to the change in the message's polarity.)

Before discussing FM any further, an important point must be made here. A squarewave message has been used in this discussion to help you visualise how an FM carrier responds to its message. In so doing, Figure 1 suggests that the resulting FM signal consists of only two sinewaves (one at a frequency above the carrier and one below). However, this isn't the case. For reasons best left to your instructor to explain, the spectral composition of the FM signal in Figure 1 is much more complex than implied.

This highlights one of the important differences between FM and the modulation schemes discussed earlier. The mathematical model of an FM signal predicts that even for a simple sinusoidal message, the result is a signal that potentially contains many sinewaves. In contrast, for the same sinusoidal message, an AM signal would consist of three sinewaves, a DSBSC signal would consist of two and an SSBSC signal would consist of only one. This doesn't automatically mean that the bandwidth of FM signals is wider than AM, DSBSC and SSBSC signals (for the same message signal). However, in the practical implementation of FM communications, it usually is.

Finally, when reading about the operation of an FM modulator you may have recognised that there is a module on the Emona Telecoms-Trainer 101 that operates in the same way - the VCO module. In fact a voltage-controlled oscillator is sometimes used for FM modulation (though there are other methods with advantages over the VCO).

The experiment

In this experiment you'll generate real FM signals using the VCO module on the Emona Telecoms-Trainer 101. First, you'll set up a model digital data signal and convert it to a line code suitable for generating Frequency Shift Keying (FSK). Next, you'll set up the VCO module to output an unmodulated carrier at a known frequency. Then you'll observe the effect of frequency modulating the VCO's output with a digital data signal and then a speech signal.

It should take you about 45 minutes to complete this experiment.

 © 2008 Emona Instruments

Equipment

- Emona Telecoms-Trainer 101 (plus power-pack)
- Dual channel 20MHz oscilloscope
- two Emona Telecoms-Trainer 101 oscilloscope leads
- assorted Emona Telecoms-Trainer 101 patch leads

Procedure

Part A – Setting up a model digital data signal and line code conversion
As you know, digital data signals switch back and forth between logic-0 and logic-1. To model a digital data signal for generating a Frequency Shift Keying (FSK), we could use a squarewave or a rectangular wave. While these signals would be adequate, it's possible to model a more realistic digital data signal using the ETT-101 and so you'll set that up first and convert it to a line code for further realism.

1. Gather a set of the equipment listed above.

2. Connect the set-up shown in Figure 2 below using the PicoScope for the scope.

   Note: Insert the black plugs of the oscilloscope leads into a ground (*GND*) socket.

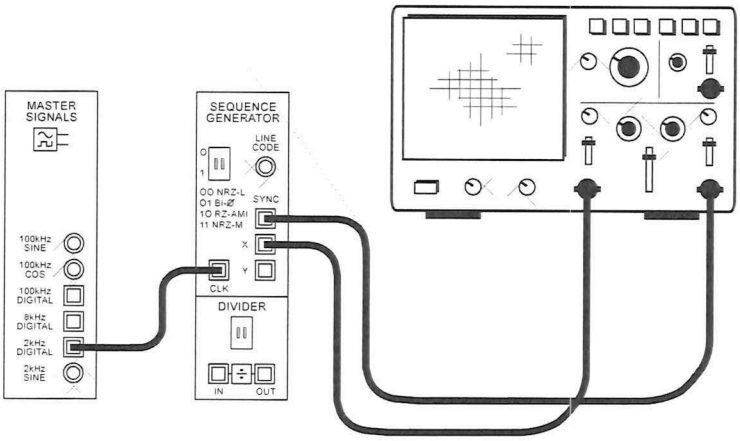


Figure 2

This set-up uses the Master Signals module's *2kHz DIGITAL* output to provide the Sequence Generator module's bit-clock. The Sequence Generator module's *X* output is a continuous 31-bit sequential binary number that's being used to model digital data using TTL logic levels. The module's *SYNC* output is a pulse that corresponds with the first bit in the *X*-output's sequence on every repetition and is provided on this module to readily facilitate scope triggering. Triggering is achieved in this experiment by connecting the module's *SYNC* output to the scope's Channel B and triggering on that input.

3. Launch the PicoScope software and press *Auto Setup*.

4. Adjust the scope's settings as followings:

   - Vertical scale for Channel A to *±10V*
   - Channel B on with its vertical scale set to *±10V*
   - Horizontal scale to *5ms*
   - Trigger Mode to *Auto*
   - Trigger Channel to *B*
   - Trigger Threshold to approx *2.5V* (if it's not already)
   - Vertical positions of the two channels so that they're not overlayed

   Note: Once done, you'll be viewing several instances of the *X* output's full 31-bit sequence digital data signal together with the module's SYNC output signal.

5. Use the scope's measurement cursors to establish that the Sequence Generator module's X-output has TTL voltage levels.

Question 1
What proves that the Sequence Generator module's X-output is a TTL signal? **Tip:** It's **not** its peak-to-peak voltage.

_____

_____

Question 2
Do you think this signal is unipolar or bipolar? Explain your answer.

_____

_____

 © 2008 Emona Instruments

Ask the instructor to check your work before continuing.

Raw digital signals like this are rarely used as the message signal in communications (for a number of reasons not discussed here). Instead, they're usually conditioned using one of a number of line codes. For this experiment, you'll use a simple line code called NRZ-L that turns the Sequence Generator module's X-output to a bipolar signal. This is ideal for the generation of a simple FSK signal using a VCO that switches between two frequencies, one above and one below its rest frequency.

6. Set the Sequence Generator module's DIP-switches to 00.

   Note: This selects NRZ-L as the line code for the Sequence Generator module's *LINE CODE* output.

7. Modify the setup as shown in Figure 3 below.

   Note: Once done, you'll be viewing the Sequence Generator module's *LINE CODE* output which is a bipolar version of its *X* output.

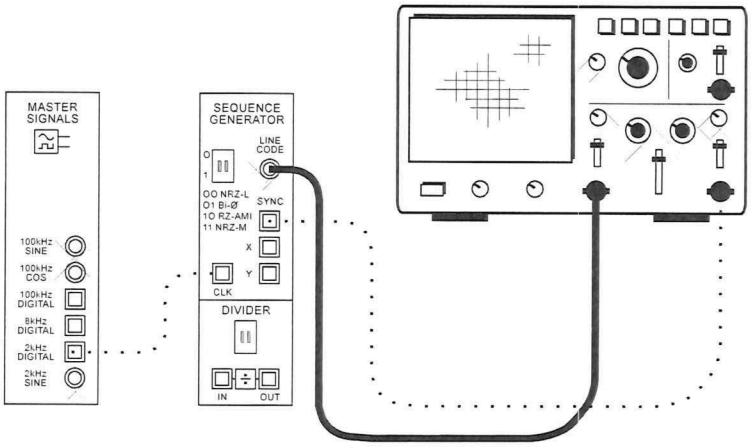


Figure 3

 © 2008 Emona Instruments

8. Use the scope's cursors to measure the line code's peak voltages relative to the baseline (ie relative to 0V).

Question 3
What proves that the Sequence Generator module's *LINE CODE* output is non-longer unipolar?

_____

_____

Ask the instructor to check your work before continuing.

 © 2008 Emona Instruments

Part B – Implementing frequency shift keying
A VCO can be used to generate an FSK signal and you'll do that next.

9. Locate the VCO module and turn its *GAIN* control fully anti-clockwise.

   Note: This turns the VCO module's sensitivity down to zero so it will ignore voltages connected to its input. You'll change this in a later step.

10. Set VCO module's frequency range to *LO*.

    Note: This sets the VCO module's output to frequencies from about 1kHz to about 17kHz.

11. Modify the setup as shown in Figure 4 below.

    Note: Once done, the scope will likely have lost triggering on Channel A but that's ok - Ignore this for now.

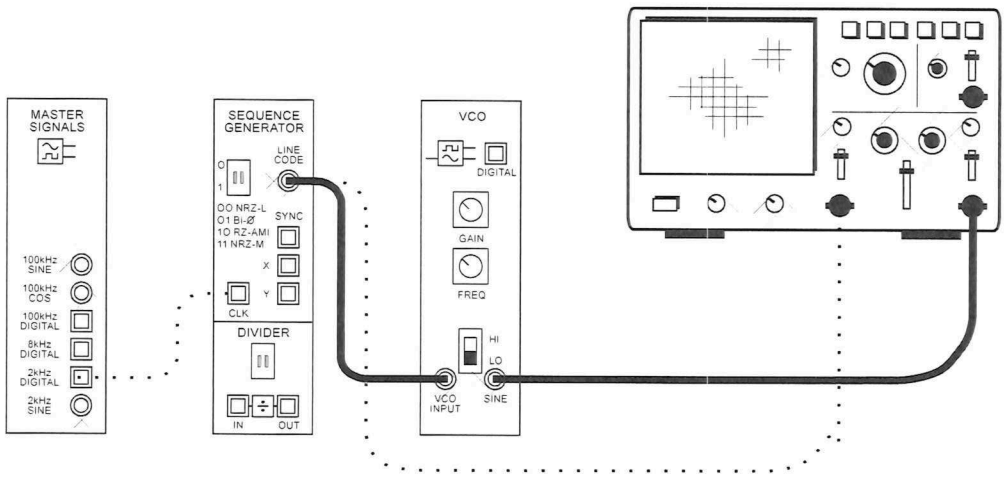


Figure 4

12. Set the scope's horizontal scale to *1ms* and its Trigger Level to *0V*.

13. Use the scope's measurement tools to help set the VCO module's output to 10kHz by adjusting the *FREQ* control.

14. Turn the VCO module's *GAIN* control to about the 12 o'clock position.

 © 2008 Emona Instruments

15. Suspend the scope's acquisition by pressing its stop button.

    Note 1: This will freeze the display so that you can look at the two signals.

    Note 2: Look at the VCO module's output signal - You should see that its frequency changes indicating that his is an FSK (FM) signal.

16. Examine the FSK signal and note where in the digital data signal that it changes frequency and in what way it changes.

Question 4
Which one of the following statements is true about the FSK signal that you've generated?

- ☐ The FSK signal's output frequency is low for logic-0 and high for logic 1
- ☐ The FSK signal's output frequency is high for logic-0 and low for logic 1

17. Restart the scope and change the scope's horizontal scale to *200μs/div.*

18. Suspend the scope's acquisition again.

19. Use the horizontal cursors to measure the period of the FSK signal when the data is logic-0. Record this in Table 1 below.

20. Use this measurement of the FSK signal's period to calculate and record its frequency.

21. Use the horizontal cursors to measure and record the period of the FSK signal when the data is logic-1. Record this in Table 1 below.

22. Repeat Step 20.

Table 1

Data's logic state	Period	Frequency
Logic-0		
Logic-1		

 © 2008 Emona Instruments

Question 5
Why was it necessary to measure the frequencies of the FSK signal by using the cursors to measure their periods? (Putting this question another way, why couldn't we use the scope's measurement tools to perform these frequency measurements?)

Question 6
Knowing the bipolar data signal's peak voltages and the FSK signal's two frequencies, it's possible to determine the VCO module's sensitivity across its range of deviation. Devise a suitable method of doing so and record your answer below, together with your working.

Question 7
What's another name for the "sensitivity" of an FM modulator? **Tip:** If you don't know, the answer is in Section 5 of the theory notes.

Question 8
In terms of its spectral composition, how many sinewaves make up the FSK signal?

 © 2008 Emona Instruments

23. Restart the PicoScope and set it to Spectrum Mode.

24. Turn off Channel A and turn on Channel B.

25. Set the Spectrum Range to 50kHz then examine the FSK signal's spectral composition. Make changes to your answer to Question 8 as necessary.

Ask the instructor to check your work before continuing.

 © 2008 Emona Instruments

Part C – Generating an FM signal using an analog signal (speech)
The next part of the experiment lets you see what an FM signal looks like when modulated by speech.

26. Return the PicoScope to Scope Mode.

27. Set the scope's horizontal scale to *1ms*.

28. Set the scope's Channel B vertical scale to *±5V*.

29. Disconnect the plugs to the Sequence Generator module's *LINE CODE* output.

30. Connect them to the Speech module's output as shown in Figure 5 below.

    Remember: Dotted lines show leads already in place.

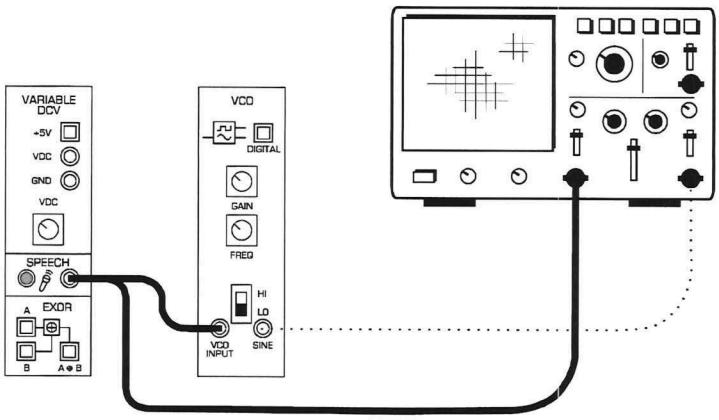


Figure 5

31. Talk or hum while watching the scope's display.

    Note: The gain of the Speech module's amplifier is quite high so you'll find that it's easy to overdrive it, causing clipping on the message which is observed on the AM signal too. Try to not talk/hum/etc too loudly to avoid this problem.

 © 2008 Emona Instruments

32. Locate the scope's stop button and hover the mouse's cursor over it (but don't click on it yet).

33. Talk or hum again, but this time, stop the scope's acquisition (to freeze the display) when your speech produces a wide variety of voltages in the message signal.

    Tip: Try varying the volume of your voice as you talk to achieve this effect (without overdriving the Speech module's amp).

Question 9
What is the relationship between the FM signal's frequency deviation (that is, the VCO module's output) and the amplitude of the message?

_____

_____

Ask the instructor to check your work before finishing.

 © 2008 Emona Instruments

Name: _____

Class: _____

15 - Amplitude shift keying

Experiment 15 - Amplitude Shift Keying

Preliminary discussion

An essential part of electronic communications and telecommunications is the ability to share the channel. This is true regardless of whether the channel is copper wire, optical fibre or free-space. If it's not shared then there can only ever be one person transmitting on it at a time. Think about the implications of this for a moment. Without the ability to share, there could only be one radio or TV station in each area. Only one mobile phone owner could use their phone in each cell at any one time. And there would only be the same number of phone calls between any two cities as the number of copper wires or optical fibres that connected them.

So sharing the channel is essential and there are several methods of doing so. One is called *time division multiplexing* (TDM) and involves giving the users exclusive access to the channel for short periods of time. On the face of it, this type of sharing might seem impractical. Imagine giving all mobile phone users in a cell just a minute or so to make their call then having to wait until their turn comes around again. However, TDM works well when the access time is extremely short (less than a second) and the rate of the sharing is fast. This allows multiple users to appear to have access all at the same time.

TDM is used for digital communications and is achieved by interleaving the users' data. That is, a portion of one user's data is transmitted followed by a portion of the next user's data and so on. Unfortunately, there's a catch. If the message is real-time information which can't afford to be delayed (like digitally encoded speech) then, as the number of users increases, so must the data's bit-rate. Experiment 14 has shown that doing so increases the likelihood of the channel's bandwidth distorting the signal causing errors at the receiver.

Another method of sharing the channel is called *frequency division multiplexing* (FDM) and involves giving the users exclusive and uninterrupted access to a portion of the channel's radio frequency spectrum. To transmit their message the user must superimpose it onto a carrier that sits inside their allocated band of frequencies. This method is used by broadcast radio and television to share free-space.

FDM is also used for digital communications and uses the same modulation schemes available to analog communications including: AM, DSBSC and FM. When AM is used for multiplexing digital data, it is known as *amplitude shift keying* (ASK). Other names include: *on-off keying, continuous wave* and *interrupted continuous wave.*

Figure 1 below shows what an ASK signal looks like time-coincident with the digital signal that has been used to generate it.

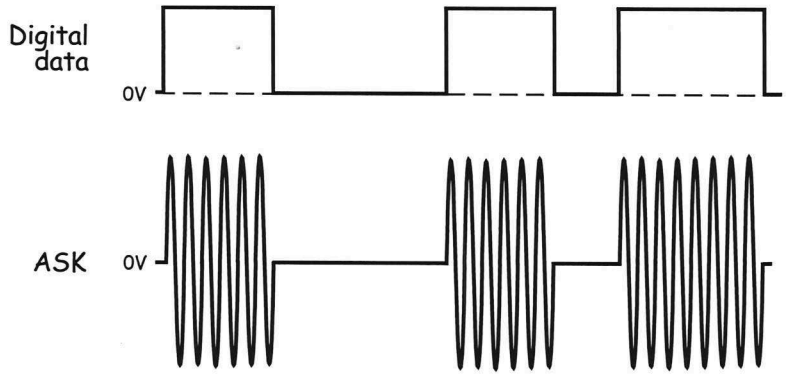


Figure 1

Notice that the ASK signal's upper and lower limits (the *envelopes*) are the same shape as the data stream (though the lower envelope is inverted). This is simultaneously an advantage and a disadvantage of ASK. Recovery of the data stream can be implemented using a simple envelope detector (refer to the preliminary discussion of Experiment 6 for an explanation of the envelope detector's operation). However, noise on the channel can change the envelopes' shape enough for the receiver to interpret the logic levels incorrectly causing errors (analog AM communications have the same problem and the errors are heard as a hiss, crackles and pops).

ASK can be generated by conventional means like the one modelled in Experiment 4. Here you'll examine the operation of an alternative method that involves using the digital signal to switch the carrier's connection to the channel on and off.

The experiment

In this experiment you'll use the Emona Telecoms-Trainer 101 to generate an ASK signal using the switching method. Digital data for the message is modelled by the Sequence Generator module. You'll then recover the data using a simple envelope detector and observe its distortion. Finally, you'll use a comparator to restore the data.

It should take you about 40 minutes to complete this experiment.

Equipment

- Emona Telecoms-Trainer 101 (plus power-pack)
- Dual channel 20MHz oscilloscope
- three Emona Telecoms-Trainer 101 oscilloscope leads
- assorted Emona Telecoms-Trainer 101 patch leads

Procedure

Part A - Generating an ASK signal

1. Gather a set of the equipment listed above.

2. Set up the scope per the instructions in Experiment 1.

3. Set the scope's *Trigger Source* control to the *EXT* position.

4. Set the scope's *Trigger Source Coupling* control to the *HF REJ* position.

5. Set the scope's Channel 1 and Channel 2 *Input Coupling* controls to the *DC* position.

6. Set the scope's *Timebase* control to the *1ms/div* position.

7. Connect the set-up shown in Figure 2 below.

   Note: Insert the black plugs of the oscilloscope leads into a ground (*GND*) socket.

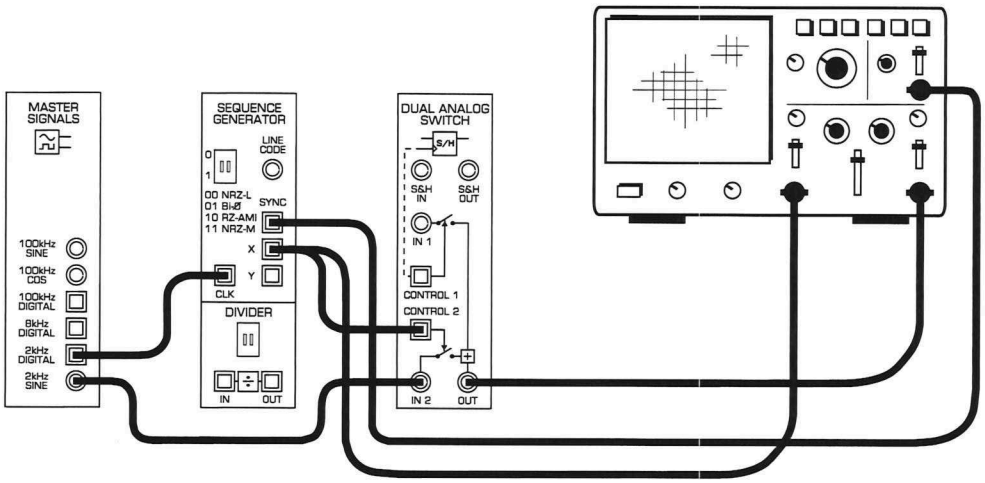


Figure 2

The set-up in Figure 2 can be represented by the block diagram in Figure 3 below. The Sequence Generator module is used to model a digital signal and its SYNC output is used to trigger the scope to provide a stable display. The Dual Analog Switch module is used to generate the ASK signal.

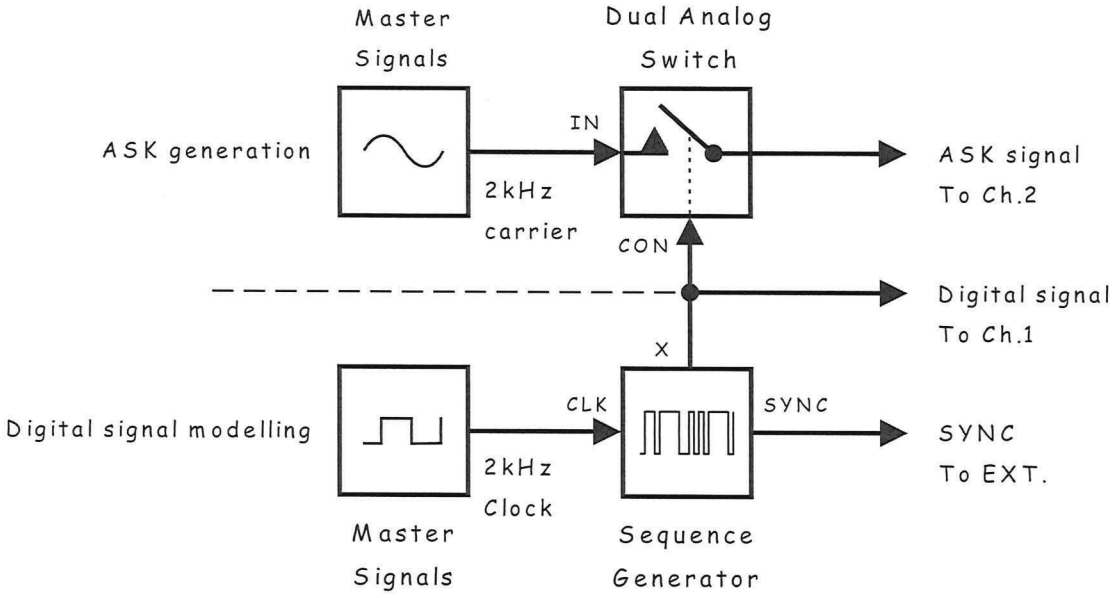


Figure 3

8. Set the scope's *Mode* control to the *DUAL* position to view the Sequence Generator module's output and the ASK signal out of the Dual Analog Switch module.

9. Compare the signals.

Question 1
What is the relationship between the digital signal and the presence of the carrier in the ASK signal?

_____

_____

Question 2
What is the ASK signal's voltage when the digital signal is logic-0? **Tip:** If you're not sure, briefly set the Channel 2 *Input Coupling* control to the *GND* position.

Ask the instructor to check your work before continuing.

Notice that the ASK signal's carrier and the Sequence Generator module's clock are the same frequency (2kHz). Moreover, notice that they're from the same source – the Master Signals module.

This has been done to make the ASK signal easy to look at on the scope. However, it makes the set-up impractical as a real ASK communications system because the carrier and the data signal's fundamental are too close together in frequency. For reasons explained in Experiment 6 (see pages 6-10 and 6-11), this makes recovering the digital data at the receiver difficult if not impossible.

Ideally, the carrier frequency should be much higher than the bit-rate of the digital signal (which is determined by the Sequence Generator module's clock frequency in this set-up). The next part of the experiment gets you to set the carrier to a more appropriate frequency (about 100kHz). In the process, the Dual Analog Switch module's output will look more like a conventional ASK signal.

10. Locate the VCO module and set its *Frequency Adjust* control to about the middle of its travel.

11. Set the VCO module's *Range* control to the *HI* position.

12. Modify the set-up as shown in Figure 4 on the next page.

    Remember: Dotted lines show leads already in place.

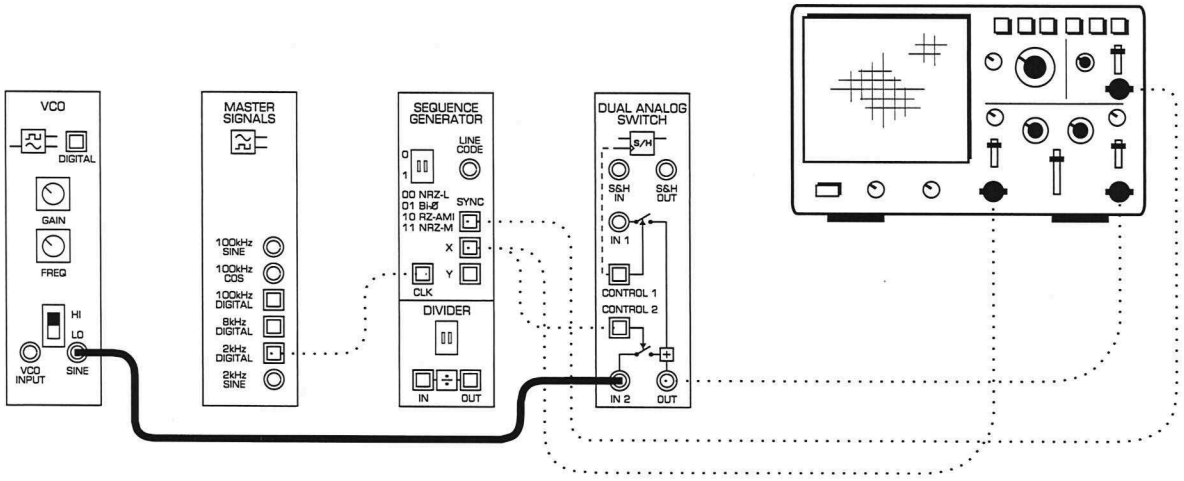


Figure 4

This set-up can be represented by the block diagram in Figure 5 below.

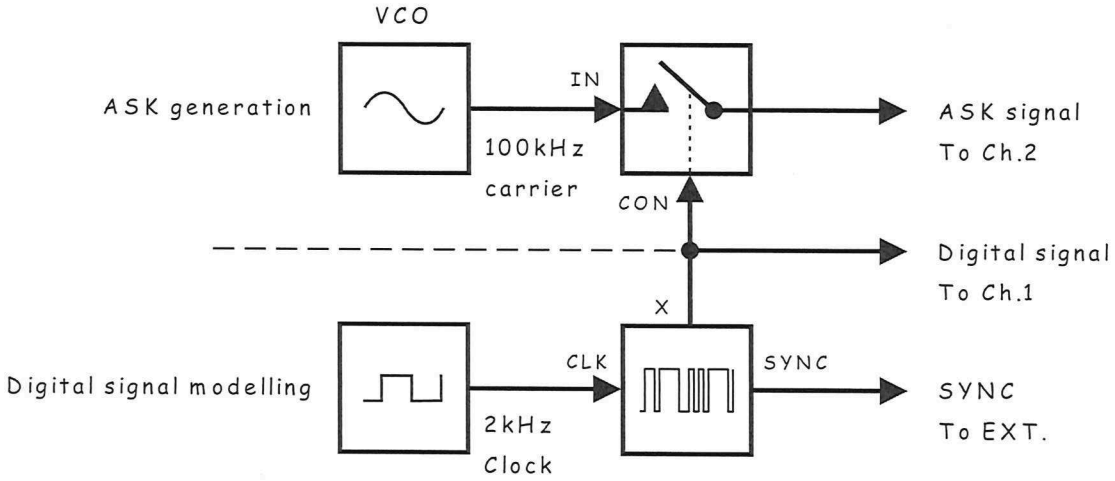


Figure 5

13. Compare the signals.

14. Use the scope's Channel 1 *Vertical Position* control to overlay the digital signal with the ASK signal's envelopes and compare them.

Question 3
What feature of the ASK signal suggests that it's an AM signal? **Tip:** If you're not sure, see the preliminary discussion.

_____

_____

Ask the instructor to check your work before continuing.

Note: Check with the teacher before attempting Part B

Part B – Demodulating an ASK signal using an envelope detector
As ASK is really just AM (with a digital message instead of speech or music), it can be recovered using any of AM demodulation schemes. The next part of the experiment lets you do so using an envelope detector.

15. Locate the Tuneable Low-pass Filter module and turn its *Gain* control fully clockwise.

16. Turn the Tuneable Low-pass Filter module's *Cut-off Frequency Adjust* control fully clockwise.

17. Modify the set-up as shown in Figure 6 below.

    Note: The left most modules have been left off to fit the drawing on the page.

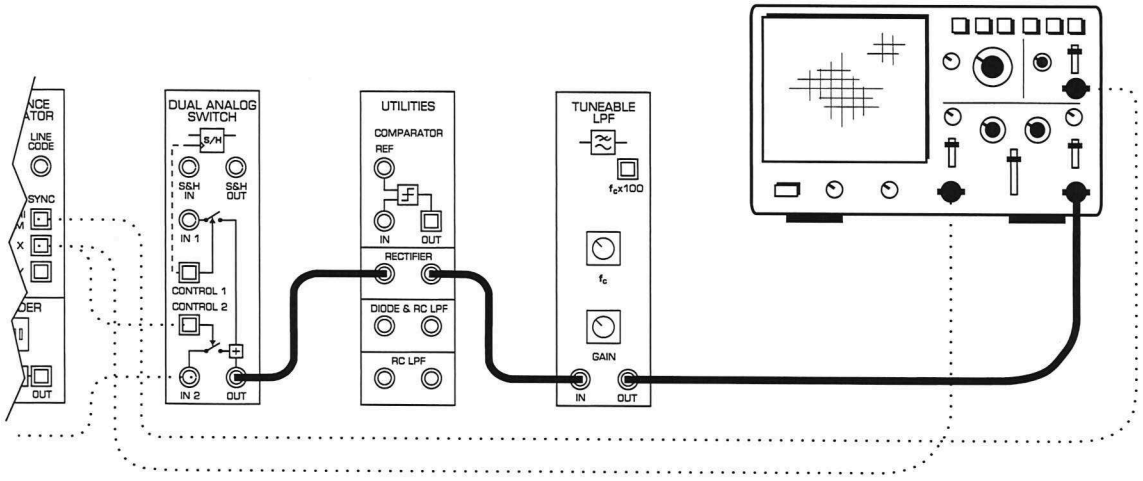


Figure 6

The ASK generation and demodulation parts of the set-up can be represented by the block diagram in Figure 7 on the next page. The rectifier on the Utilities module and the Tuneable Low-pass filter module are used to implement an envelope detector to recover the digital data from the ASK signal.

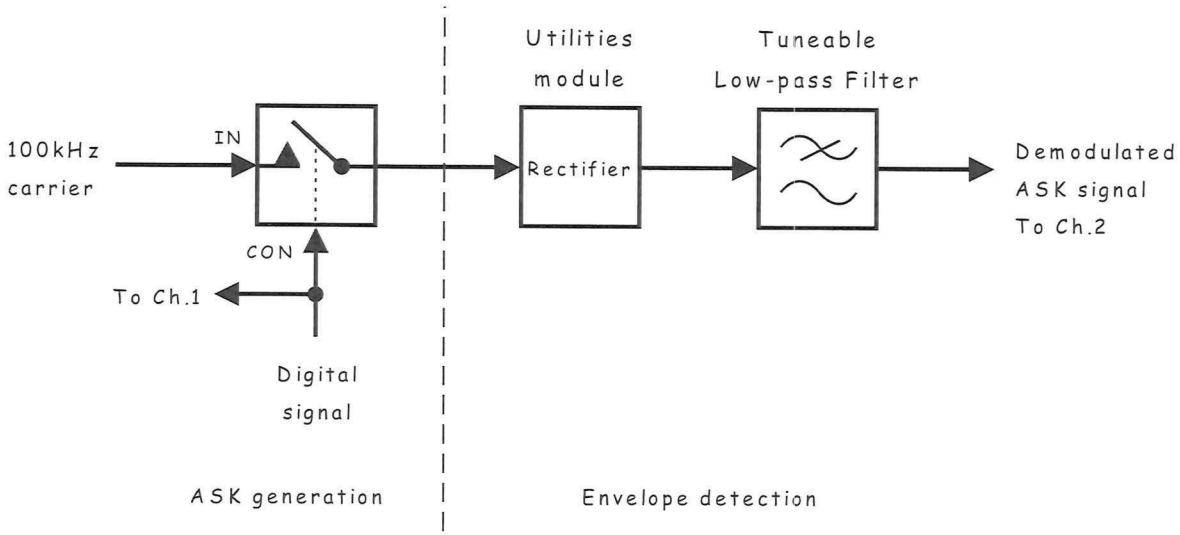


Figure 7

18. Compare the original and recovered digital signals.

Question 4
Why is the recovered digital signal not a perfect copy of the original?

_____

_____

Question 5
What can be used to "clean-up" the recovered digital signal?

_____

Ask the instructor to check your work before continuing.

Part C – Restoring the recovered digital signal using a comparator

Experiment 14 shows that the comparator is a useful circuit for restoring distorted digital signals. The next part of the experiment lets you use a comparator to clean-up the demodulated ASK signal.

19. Modify the set-up as shown in Figure 8 below.

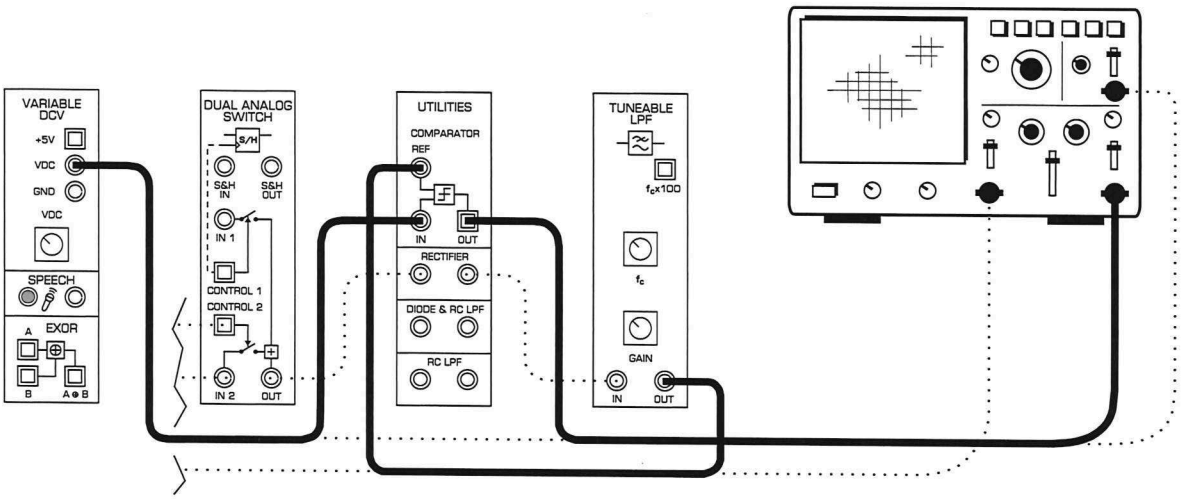


Figure 8

The ASK generation, demodulation and digital signal restoration parts of the set-up can be represented by the block diagram in Figure 9 below.

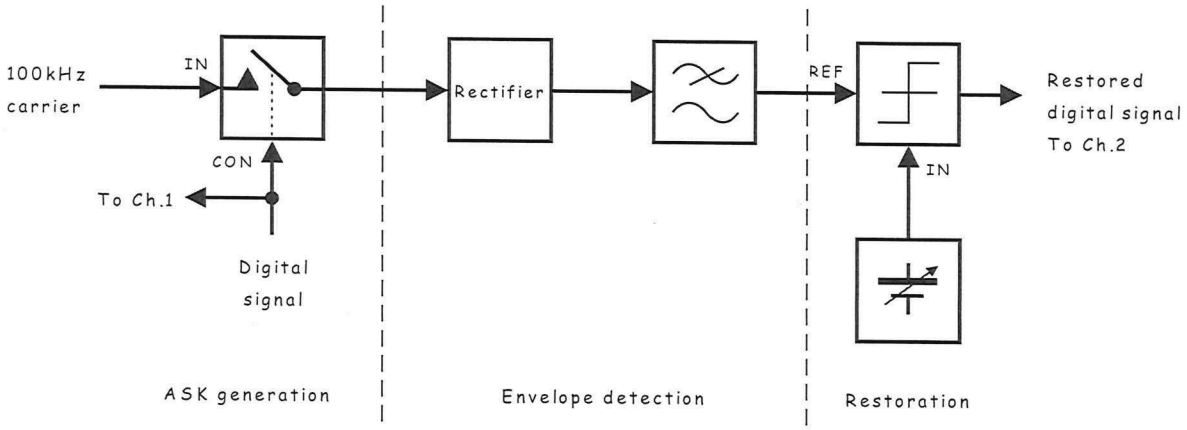


Figure 9

20. Set the Variable DCV module's *Variable DC* control to about the middle of its travel.

21. Compare the signals. If they're not the same, vary the Variable DCV module's *Variable DC* control until they are.

Question 6
How does the comparator turn the slow rising voltages of the recovered digital signal into sharp transitions?

_____

_____

Ask the instructor to check your work before finishing.

Noise

It's common for radio frequency communications systems to be upset by unwanted electromagnetic radiation called *noise*. Some of this radiation occurs naturally and is generated by the Sun and atmospheric activity such as lightning. Much of the radiation is also human-made - either unintentionally (the electromagnetic radiation given off by electrical machines and electronics equipment) or intentionally (other peoples' communication transmissions that we don't want to receive).

Most noise gets added to signals while they're in the channel. This changes the signals' shape which in turn changes how the signal sounds when demodulated by the receiver. If the noise is sufficiently big (relative to the size of the signal) the signal can be changed so much that it cannot be demodulated.

It's possible to model noise being added to a signal in the channel of a communications system using the Emona Telecoms-Trainer 101. If the instructor allows, this activity gets you to do so.

1. Connect the set-up shown in Figure 1 below but **don't** disconnect any of your existing wiring.

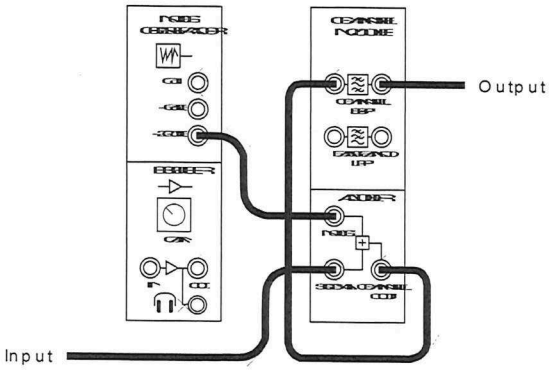


Figure 1

This set-up can be represented by the block diagram in Figure 2 on the next page. It models the behaviour of a real channel by adding noise to communications signals such as ASK.

Usefully, the amount of noise can be varied by selecting either the *-20dB* output (noise is about one-tenth the size of the signal), the *-6dB* output (noise is about half the size of the signal) or the *0dB* output (noise is about the same size as the signal).

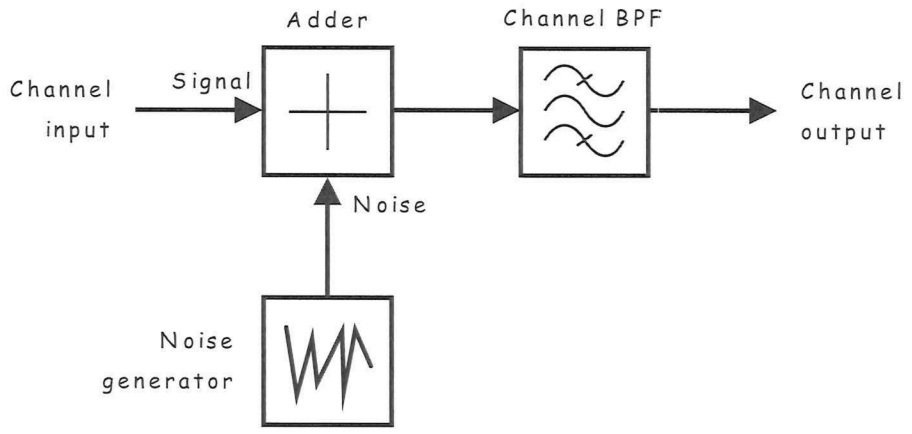


Figure 2

2. Unplug the patch lead to the Dual Analog Switch module's output and connect the noisy channel's input to it.

3. Unplug the patch lead to the rectifier's input and connect the noisy channel's output to it.

   Note: Once done, the transmitter's signal (the Dual Analog Switch module's output) travels to the receiver's input (the rectifier's input) via the model of a noisy channel.

4. Compare the original and recovered data. If they're not the same, vary the Variable DCV module's *Variable DC* control until they are.

5. Unplug the scope's Channel-2 input from the comparator's output and connect it to the Adder module's output to observe the noisy ASK signal.

6. Connect the Adder module's *Noise* input to the Noise Generator module's -6dB output to increase the noise in the channel.

7. Observe the effect that this has on the ASK signal.

8. Reconnect the scope's Channel-2 input to the comparator's output.

9. Compare the original and recovered data. If they're not the same, vary the Variable DCV module's Variable DC control until they are.

   Note: It may be impossible to recover the data.

Name: _____

Class: _____

4 - FM demodulation using the Phase-locked Loop

Experiment 4 - FM demodulation using the Phase-locked Loop

Preliminary discussion

There are several methods of demodulating an FM signal. An earlier experiment introduced you to the zero-crossing detector method. This experiment demonstrates the more widely used *phase-locked loop* (PLL) method. The block diagram of a PLL used for FM demodulation is shown in Figure 1 below.

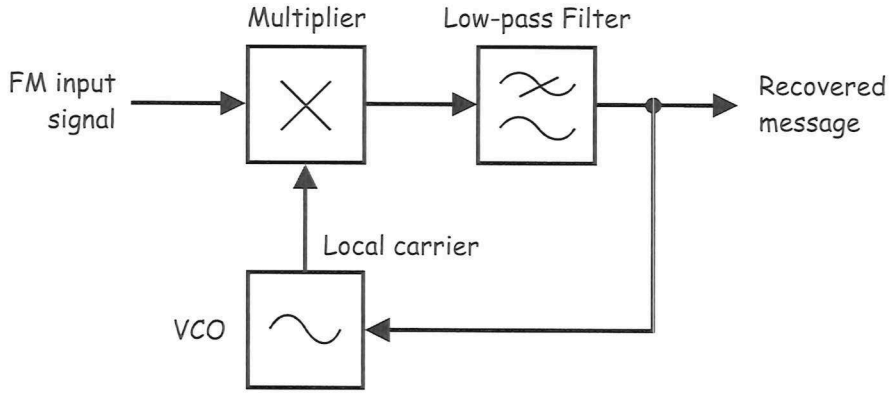


Figure 1

As you can see, a VCO is used to generate a local carrier that is adjusted for an output frequency that is reasonably close to the carrier frequency of the FM signal to be demodulated. As a mathematical explanation of PLL operation is complicated, it's simpler to think of its operation by first imagining that the connection between the low-pass filter and the VCO is broken. If this is the case, the arrangement is basically the product detector shown in Figure 2 below.

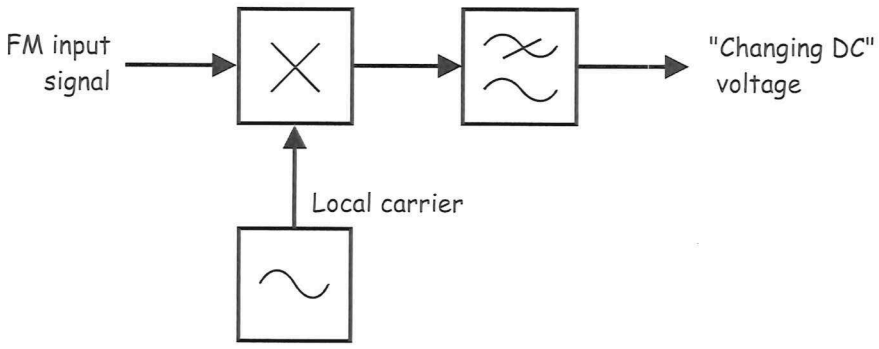


Figure 2

Suppose the FM input signal is unmodulated and the VCO's output is a sinewave with an identical frequency. The math predicts that the multiplier's output is a DC component and a sinewave equal to the sum of the inputs' frequencies. Importantly, the magnitude of the DC component is a function of the phase difference between the two carriers. A phase difference of 0° produces the maximum DC output voltage and a phase difference of ±90° produces 0V DC. Put another way, any phase difference between the multiplier's inputs either side of 90° produces a DC voltage.

Now, suppose that the VCO's output changes to a slightly different frequency (which it's likely to be in the first instance). This has two effects on the multiplier's output. First, the frequency of the sinewave on its output changes. Second and more importantly, the resulting continuous change in phase difference between the carriers causes the magnitude of the DC component to change continuously in sympathy (so it's not really DC any more).

The low-pass filter removes the multiplier's sinewave but let the "changing DC" pass. As the low-pass filter's output is connected to the VCO's input (per Figure 1), this changing DC component causes the VCO's output frequency to change as well. Inevitably, at some point the unmodulated FM input signal and the VCO's output will have exactly the same frequency with a 90° phase shift. When this happens, the multiplier's DC component is 0V and the PLL achieves "lock".

Now, once lock is achieved, any changes in the frequency of the FM input signal automatically produce a DC voltage on the output of the low-pass filter that forces the VCO's output frequency to track it. Importantly, as frequency changes in the FM input signal are caused by a message at the modulator, the changing DC voltage on the low-pass filter's output must be a copy of the message.

The experiment

In this experiment you'll use the Emona Telecoms-Trainer 101 to generate an FM signal using a VCO. Then you'll set-up a phase-locked loop (PLL) using a second VCO module on the ETT-101-21 expansion board and use it to demodulate the FM signal.

It should take you about 40 minutes to complete this experiment.

Pre-requisites:

Experiment 1 (Vol. 1): Setting up an oscilloscope
Experiment 2 (Vol. 1): An introduction to the Telecoms-Trainer 101
Experiment 3 (Vol. 1): Modelling equations
Experiment 5 (Vol. 1): DSBSC modulation
Experiment 7 (Vol. 1): DSBSC demodulation
Experiment 1 (Vol. 2): AM (method 2) and product detection of AM signals (desirable)

Equipment

- Emona Telecoms-Trainer 101 (plus power-pack)
- Desktop/laptop with PicoScope software installed
- Emona ETT-101-21 LCD & PLL board
- Two Emona Telecoms-Trainer 101 oscilloscope leads
- Assorted Emona Telecoms-Trainer 101 patch leads

Procedure

Part A – Setting up the FM modulator

To experiment with FM demodulation you need an FM signal. The first part of the experiment gets you to set one up.

1. Gather a set of the equipment listed above.
2. **Turn off** the Emona Telecoms-Trainer.
3. Insert the ETT-101-21 board into the Emona Telecoms-Trainer's *EXPANSION* socket.
4. Turn the Emona Telecoms-Trainer back on.
5. Locate the ETT-101's VCO module (**not** the VCO module on the expansion board) and set its *Range* control to the *HI* position.
6. Set the VCO module's *Gain* control to about the middle of its travel.

7. Connect the set-up shown in Figure 3 below.

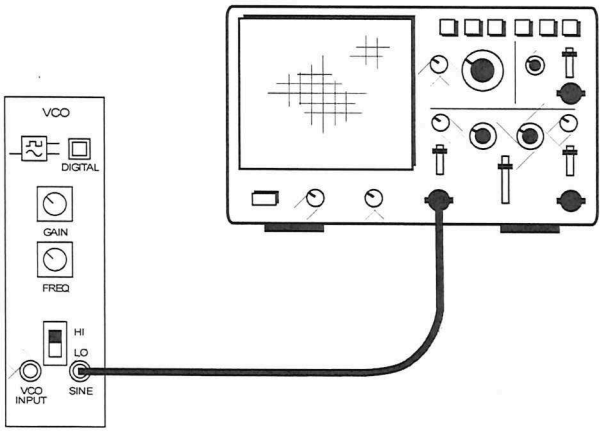


Figure 3

8. Launch the scope's app and click on *Auto Setup*.

   Note: The scope's display won't be triggered but that's ok for now.

9. Adjust the following scope settings as indicated:

   - Channel A Scale: *±5V*
   - Channel A Input Coupling: *DC*
   - Timebase: *5μs/div*
   - Trigger Mode: Auto
   - Trigger Source: A
   - Trigger Threshold: 0V

   Note: Some of these options are likely already on the setting specified above by default.

10. Adjust the VCO module's *Frequency Adjust* control for 100kHz on the *SINE* output.

    Tip: Use the scope's measurement tools to perform this step.

11. Connect the set-up shown in Figure 4 below.

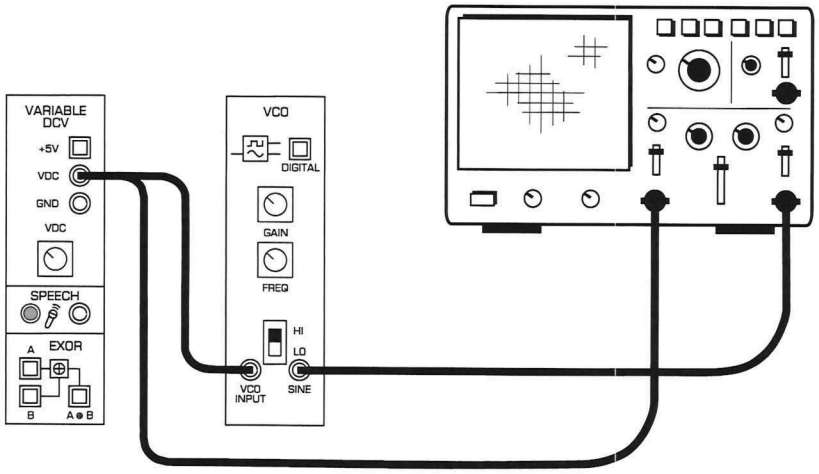


Figure 4

This set-up can be represented by the block diagram in Figure 5 below. The Variable DCV module is being used to provide a simple DC message and the VCO module implements an FM modulator with a carrier frequency of 100kHz.

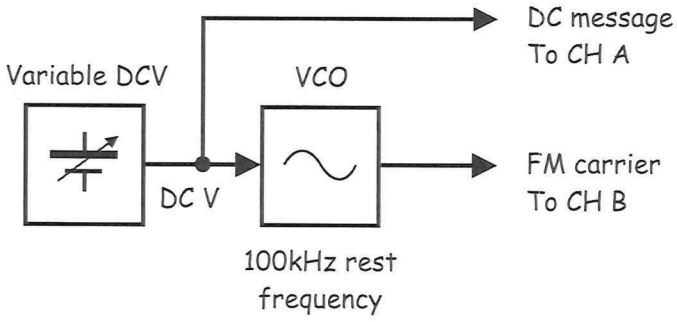


Figure 5

12. Adjust the following scope settings as indicated:

- Channel B Scale: *±5V* (this will turn on Channel B if it's not already)
- Channel A Input Coupling: *DC*
- Trigger Source: B

Note: Don't worry that the two channels are overlayed.

13. Vary the Variable DCV module's *DC Voltage* control and check that the VCO module's output frequency changes accordingly.

Ask the instructor to check your work before continuing.

For a variety of reasons, an important operating parameter of FM modulators is its *sensitivity*. This is how much the FM modulator's output frequency deviates from the carrier (or rest) frequency for a given change in input voltage. It is typically expressed in Hertz per volt (ΔHz/ΔV).

The sensitivity of the FM modulator that you've just wired will be set to 2.5kHz/Volt by adjusting the VCO module's *GAIN* control. The following instructions get you to do so.

14. Set the Variable DCV module's output to +2V.

15. Adjust the VCO module's *GAIN* control for a 105kHz output.

    Note 1: Again, use the scope's measurement tools to perform this step.

    Note 2: As a 2V DC input causes the VCO module's output frequency to change from 100kHz to 105kHz then the sensitivity must be 2.5kHz/Volt.

16. Set the Variable DCV module's output to -2V.

17. Note the VCO module's new output frequency.

    Note: Assuming the VCO module's operation is reasonably linear, the new output frequency should be about 95kHz.

Ask the instructor to check your work before continuing.

18. Disconnect the Variable DCV module and modify the set-up as shown in Figure 6 below.

Note: Leave the scope's Channel A input connected to the Variable DCV module's output for now.

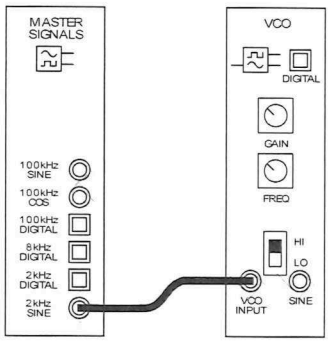


Figure 6

This set-up can be represented by the block diagram in Figure 7 below. A 2kHz sinewave is being used to model the message and the output of the VCO module is the corresponding FM signal. You should notice the FM signal jitters. This is because its changing frequency is constantly changing.

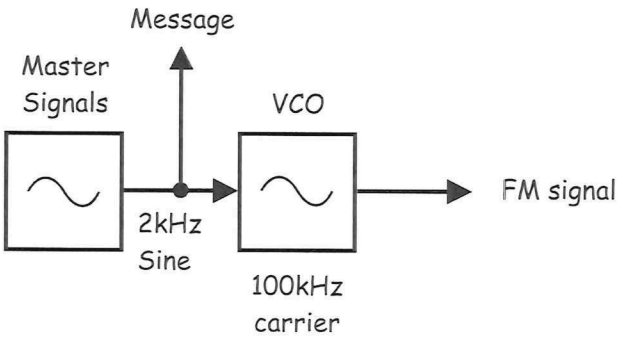


Figure 7

Ask the instructor to check your work before continuing.

Part B – Setting up the PLL demodulator

Now that you have an FM modulator, Part B gets you to set the sensitivity of the VCO module on the ETT-101-21 expansion board then use it to demodulate the FM signal. Importantly, although setting the sensitivity of the VCO module requires repeating several steps in Part A, you are not setting up another FM modulator.

19. Disconnect the scope's Channel B from the VCO module's output but **don't** dismantle the existing set-up or make any changes to its settings.

20. Locate the ETT-101-21 expansion board's VCO module and set its *Range* control to the *HI* position.

21. Set the ETT-101-21 VCO module's *Gain* control to about the middle of its travel.

The sensitivity of the FM demodulator is another important parameter in FM communications that can be optimised here. The following instructions get you to do so.

22. Connect the set-up in Figure 3 again using the ETT-101-21 expansion board's VCO module.

    Note: The display will be unstable. There's no need to change the trigger source to Channel A but, if you prefer to do so, return it to Channel B for Step 24 below.

23. Adjust the VCO module's *Frequency Adjust* control for 100kHz on the *SINE* output.

24. Connect the set-up in Figure 4 again using the ETT-101-21 expansion board's VCO module.

25. Set the Variable DCV module's output to +2V.

26. Adjust the VCO module's *Gain* control for a 106kHz output.

    Note: This will set the VCO module's sensitivity to 3kHz/Volt.

27. Set the Variable DCV module's output to -2V.

28. Check that the VCO module's output frequency is about 94kHz.

Ask the instructor to check your work before continuing.

29. Set the scope's *trigger source* control to Channel A.

30. Set the scope's *timebase* to *200µs/div.*

31. Disconnect the Variable DCV module from the expansion board's VCO module and add to the set-up as shown in Figure 8 below.

    Note: There are two VCO modules in this patch diagram so take care to use the correct one in each case.

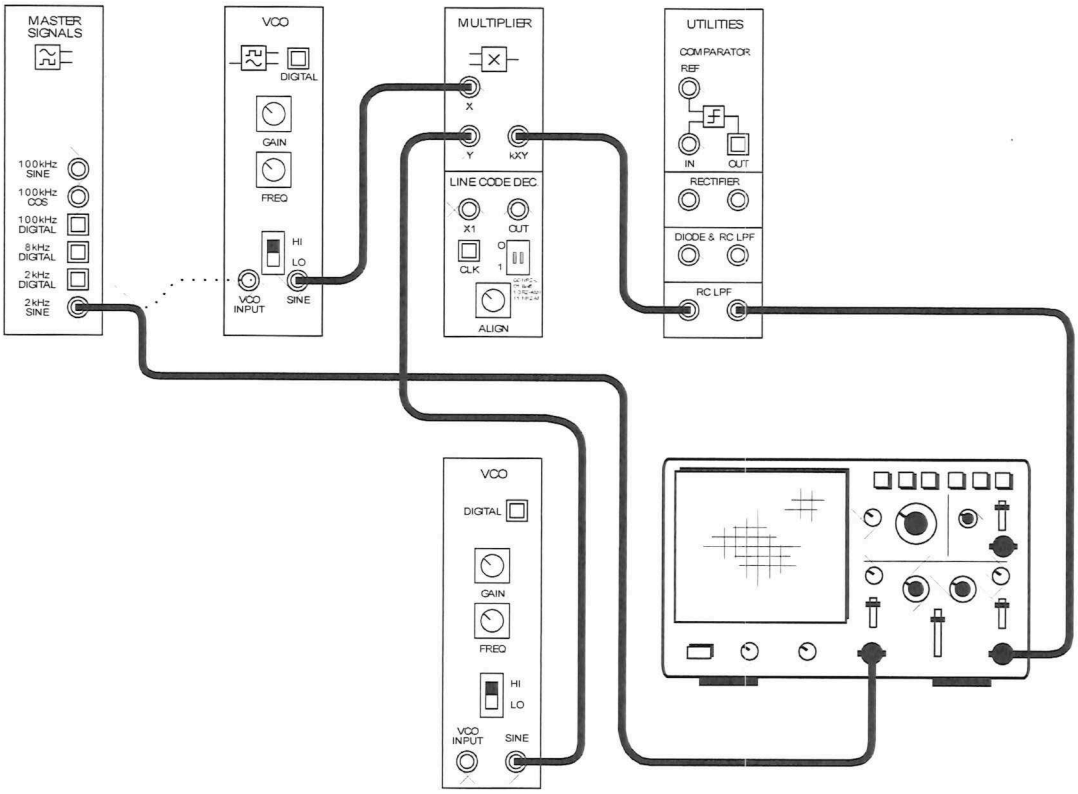


Figure 8

The additions to the set-up can be represented by the block in Figure 9 on the next page. As you can see, they implement almost all of the PLL demodulator described in the preliminary discussion.

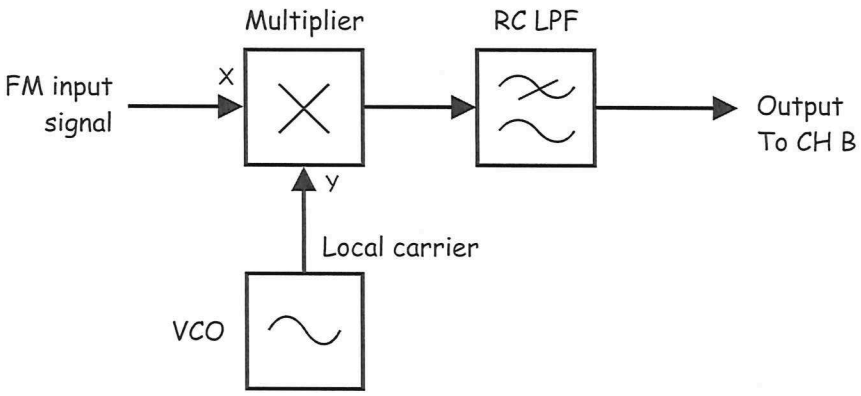


Figure 9

32. Observe the signal on the output of the RC LPF module.

Note: You'll see that the message is not being recovered.

Question 1
What small modification to the set-up in Figure 8 would complete the PLL?

_____

_____

Question 2
What would this modification enable the VCO to do that it's not currently doing?

_____

_____

Question 3
Why would this allow the original message to be recovered?

_____

_____

Ask the instructor to check your work before continuing.

33. Add patch lead shown in Figure 10 below.

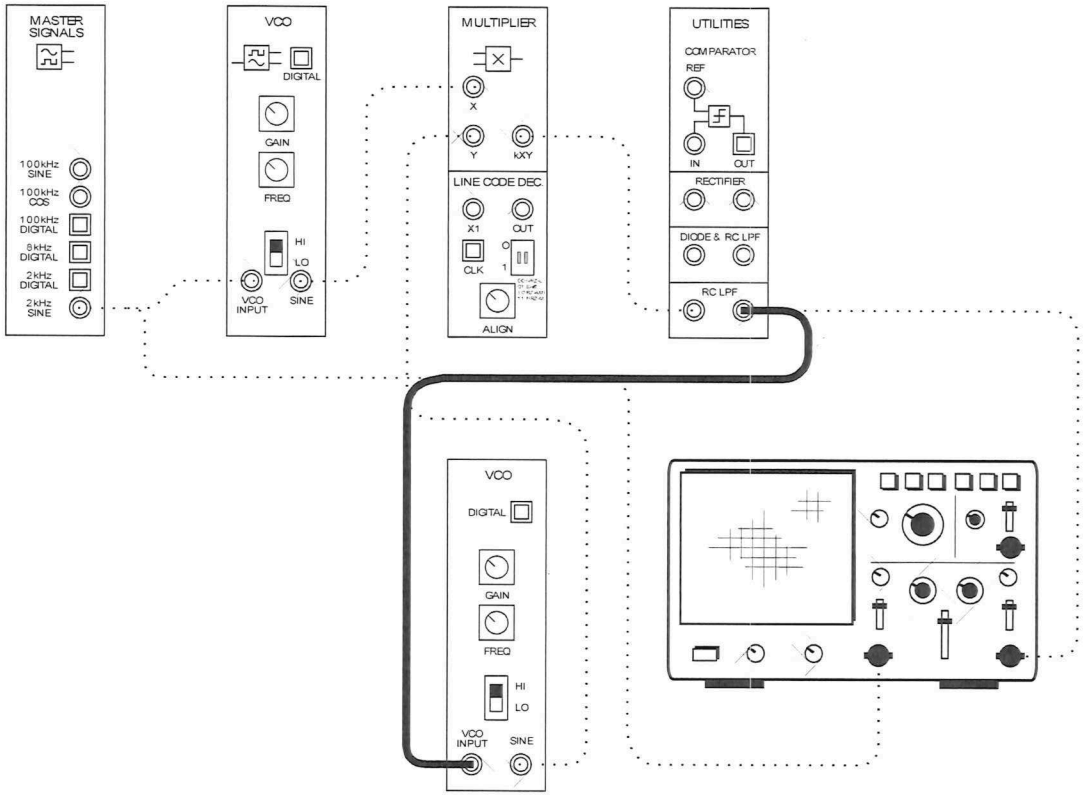


Figure 10

This final set-up can be represented by the block in Figure 11 on the next page. It implements a full PLL demodulator.

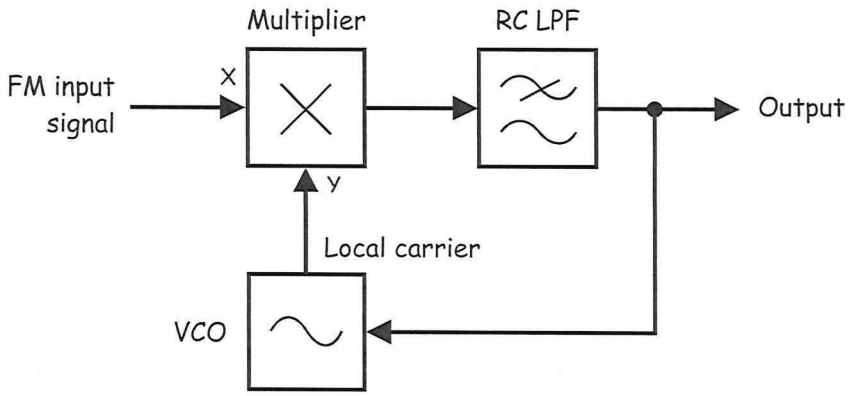


Figure 11

The entire set-up with FM modulator and demodulator can be represented by the block diagram in Figure 12 below.

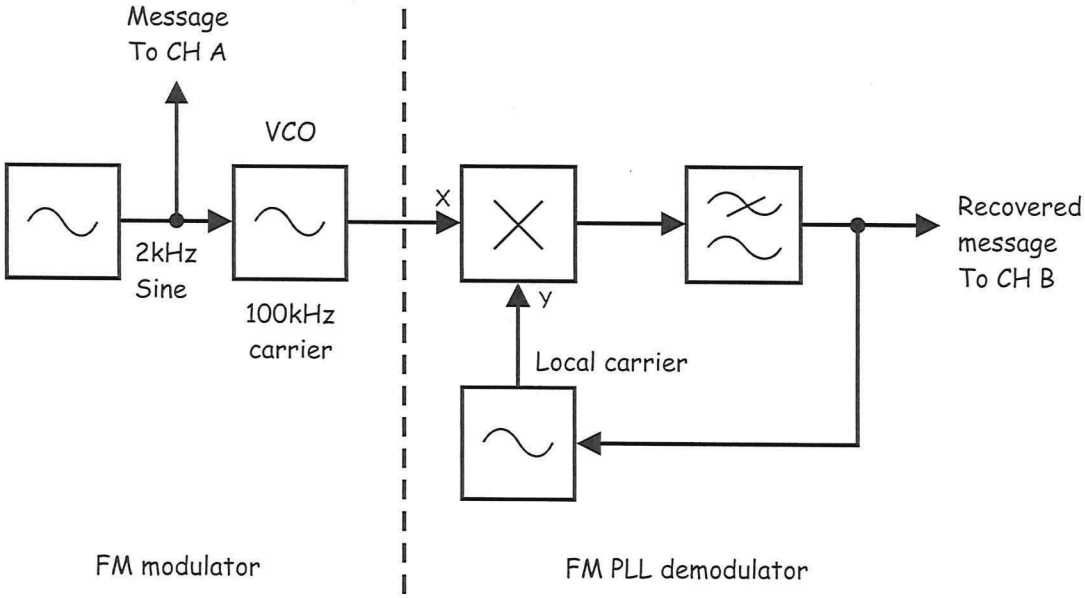


Figure 12

34. Observe the signal on the output of the RC LPF module.

    Note: You'll see that the message **is** now being recovered.

35. If the recovered message is a little distorted, make **very small** adjustments to the ETT-101-21's VCO module's *Frequency Adjust* control to obtain an undistorted message.

    Question 4
    Once the PLL is successfully recovering the message from the FM signal, what is the term that can be used to describe the PLL's condition?

    _____

To prove that the PLL is in this condition...

36. Disconnect one end of the patch lead that you installed for Step 33 then reconnect it again and observe the effects.

Ask the instructor to check your work before continuing.

Part C – Transmitting and recovering speech and music using FM

The next part of the experiment lets you use the set-up to modulate, transmit and demodulate speech and music.

37. Set the scope's timebase to *2ms/div.*

38. Disconnect the plug to the Master Signals module's *2kHz SINE* output.

39. Modify the set-up as shown in Figure 13 below.

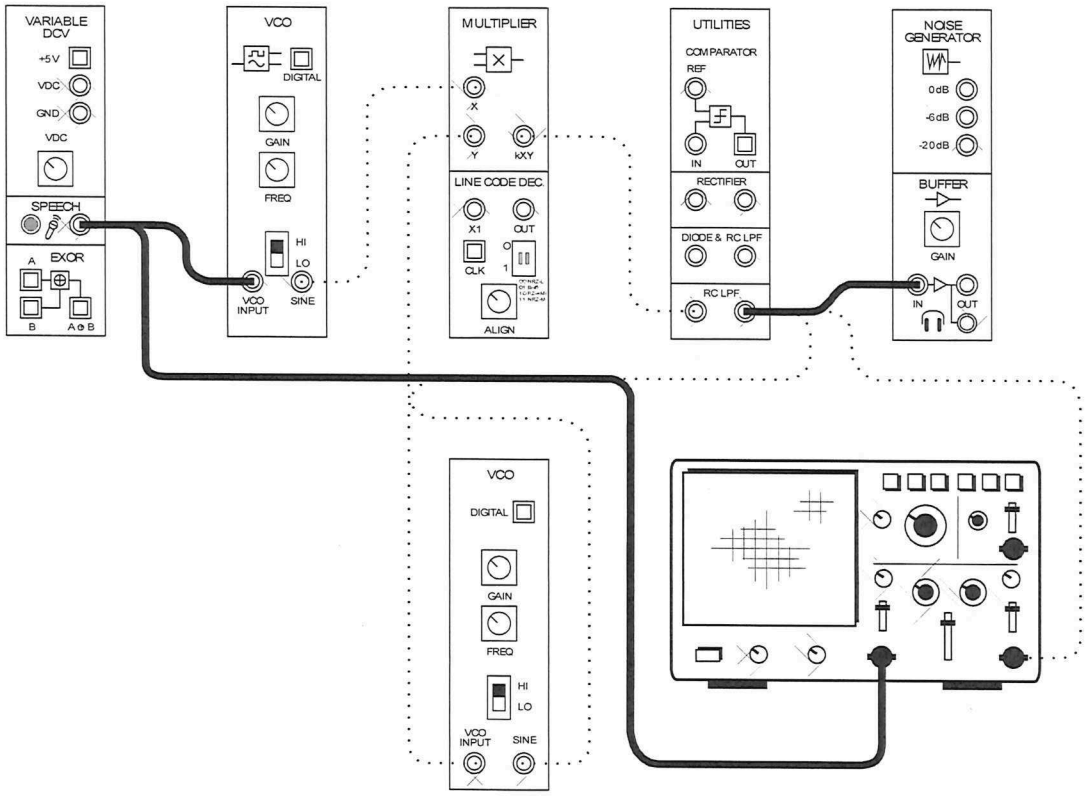


Figure 13

40. Turn the Buffer module's *Gain* control fully anti-clockwise.

41. Without wearing the headphones, plug them into the Buffer module's headphone socket.

42. Put the headphones on.

43. As you perform the next step, set the Buffer module's *Gain* control to a comfortable sound level.

44. Talk, sing or hum while watching the scope's display and listening on the headphones.

For a little fun...

45. Place an MP3 player or smartphone very close to the microphone, with music playing **very** softly.

46. Watch the scope's display and listen to the music on the headphones.

Ask the instructor to check your work before finishing.

Name: _____

Class: _____

Frequency translating AM signals using heterodyning

Experiment - Frequency translating AM signals using heterodyning

Preliminary discussion
Frequency translating any modulated signal involves reproducing an identical version of that signal but on a different carrier frequency.

For an example using amplitude modulation, suppose that a channel transmits a 5kHz tone on a 900kHz carrier using a depth of modulation of, say, 50%. The AM signal would consist of the 900kHz carrier and sidebands at 895kHz (the LSB) and 905kHz (the USB) all at certain amplitudes relative to each other. If this signal is frequency translated, say onto 455kHz, then an completely new AM signal is produced onto a 455kHz carrier and will have sidebands at 450kHz (the LSB) and 460kHz (the USB) at the same relative amplitudes.

Importantly, these numbers tell us that the frequency and size of the original message is still encoded within the new AM signal. In other words, there's no loss of information about the original message.

One method of frequency translating modulated signals (or any signal, for that matter) involves "heterodyning", better known these days as "RF mixing". This involves any process that multiplies the original signal with another sinewave. As you know, multiplying sinewaves produces sum and difference frequencies for all sinewaves involved while the original sinewaves are all theoretically lost (assuming none of the signals consist of a DC component).

Continuing the example from above, suppose that our modulated 900kHz AM signal is mixed with a 1355kHz sinewave. A new AM signal is produced consisting of sinewaves at 455kHz, 450kHz and 460kHz (ie the difference between 1355kHz and 900kHz/895kHz/905kHz) and a second AM signal is produced consisting of sinewaves at 2255kHz, 2250kHz and 2260kHz (ie the sum of 1355kHz and 900kHz/895kHz/905kHz). In the process, the sinewaves at 900kHz, 895kHz, 905kHz and 1355kHz are all theoretically lost.

The AM signal with the first set of frequencies is said to be the "down-converted" version of the original AM signal and the other is the "up-converted" version. Without any context, neither of these two new AM signals is better than the other. That said, AM receivers that use the superheterodyne for tuning always pass the down-converted version on to the demodulator to recover the message.

The experiment

In this experiment you'll use the Emona Telecoms-Trainer 101 to generate an AM signal by implementing its mathematical model. Then you'll frequency translate the AM signals onto two new carriers, one lower than the original carrier frequency (down conversion) and one higher (up conversion.

It should take you about 45 minutes to complete this experiment.

Equipment

- Emona Telecoms-Trainer 101 (plus power-pack)
- Dual channel 20MHz oscilloscope
- two Emona Telecoms-Trainer 101 oscilloscope leads
- assorted Emona Telecoms-Trainer 101 patch leads

Procedure

Part A – Setting up the AM modulator

To experiment with AM demodulation you need an AM signal. The first part of the experiment gets you to set one up.

1. Gather a set of the equipment listed above.

2. Locate the Adder module and turn its *G* and *g* controls fully anti-clockwise.

3. Locate the Variable DCV module and turn its *DC Voltage* control almost fully anti-clockwise.

4. Connect the set-up shown in Figure 1 below using the PicoScope.

   Note: Insert the black plugs of the oscilloscope leads into a ground (*GND*) socket.

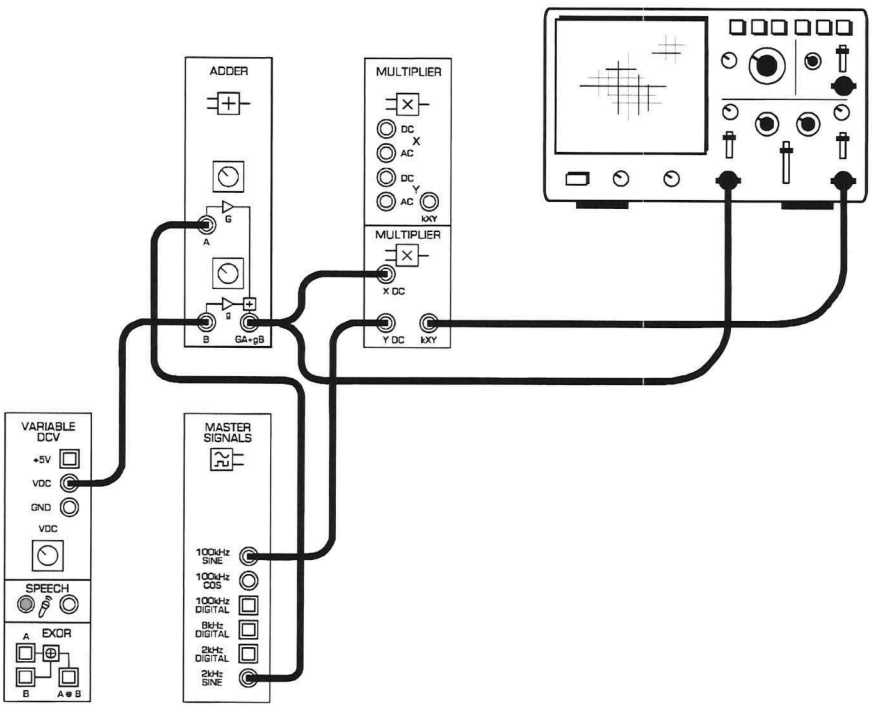


Figure 1

The set-up in Figure 2 can be represented by the block diagram in Figure 2 below. It generates a 100kHz carrier that is amplitude modulated by a 2kHz sinewave message.

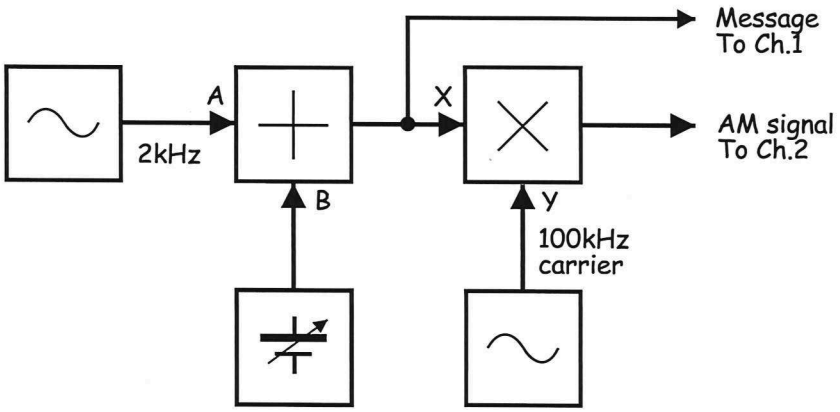


Figure 2

Note: Although the scope is now displaying the Adder module's output, it'll just be a straight line because its output is 0V.

5. Launch the PicoScope app and press *Auto Setup*.

6. Adjust the scope's settings as followings:

   - Vertical scale for Channel A to *±2V*
   - Horizontal scale to *200µs*
   - Trigger Mode to *Auto*
   - Trigger Channel to *A*
   - Trigger Threshold to *1V*

   Note: Leave Channel B turned off for the moment.

7. While watching the Adder module's output on the scope, turn its *g* control clockwise until the DC level is 1V.

   Note: You can do this by visual inspection to get it close enough (there's no need to use the scope's measurement cursors unless you would prefer to).

8. While watching the Adder module's output on the scope, turn its *G* control clockwise to obtain a 1Vp-p sinewave.

   Note: Again, you can do this by visual inspection to get it close enough.

9. Turn on the scope's Channel B and set its vertical scale to ±5V.

10. Adjust the vertical positions of the two channels so that they're not overlayed.

11. Verify that the signal on Channel B is an AM signal.

Question 1
How do you know for certain that the signal on Channel B is DSBFC AM?

_____

_____

_____

12. Call the teacher who will adjust the AM signal's depth of modulation to a random value.

13. Measure the AM signal's depth of modulation recording your measurements and calculation in Table 1 below.

Tip: If you're not sure how to do this, see Section 4 of the workbook.

Table 1

B dimension	A dimension	m

Ask the instructor to check your work before continuing.

14. Adjust the depth of modulation of your AM signal so that its modulation index is 1 (ie 100%).

Question 2
What's the benefit of using a modulation index of 1 when transmitting AM signals?

Ask the instructor to check your work before continuing.

15. Adjust the depth of modulation of your AM signal so that it's noticeably over-modulated.

Question 3
What problem occurs at the receiver when trying to demodulate this signal and why does this happen?

Ask the instructor to check your work before continuing.

16. Set the AM signal's depth of modulation to **just below 1** (ie just below 100%).

    Note: The objective with this task is to have a depth of modulation approaching 100% while ensuring that it's definitely not over-modulated.

17. Set the PicoScope to Spectrum Mode.

18. Turn off Channel A.

19. Set the Spectrum Range to 400kHz.

    Note: Once done, the spectrum will appear to have many significant sinewaves. However, this is misleading because the vertical scale is logarithmic. Most of them are actually insignificant relative to the spectral components that make up the down-converted and up-converted versions of the original AM signal. (Remember, a sinewave that is -20dB relative to another is one-tenth its size. A sinewave that is -40dB relative to another is one-hundredth its size.)

20. Identify the spectral components of the AM signal.

    Note: You may find it helpful to use one of the PicoScope's vertical cursors and align it to the AM signal's carrier (which we know to be 100kHz).

21. Increase the AM signal's depth of modulation so that its over-modulated again while watching the signal's spectral composition.

    Note: Setting the Adder module's *G* control to about the 12 o'clock position should introduce sufficient over-modulation for our purposes here.

Question 4
What happens to the AM signal's spectral composition when it is over-modulated?

_____

_____

Question 5
What's the name for this effect and what problem can it cause?

_____

_____

_____

Ask the instructor to check your work before continuing.

Part B – Frequency translation of the AM signal onto new carriers
The next part of this experiment gets you to frequency translate the AM signal onto two completely new carriers using multiplication.

22. Return the PicoScope to Scope Mode.

23. Set the AM signal's depth of modulation to about 0.15 (or 15%).

    Note: Setting the message to about 250mVpp should be close enough.

24. Locate the VCO module and turn its *Frequency Adjust* control fully clockwise.

25. Set the VCO module's *Range* control to the *HI* position.

26. Connect the scope's Channel B to the VCO module's output as shown in Figure 3 below.

    Note: Don't dismantle any other element of the previous set-up.

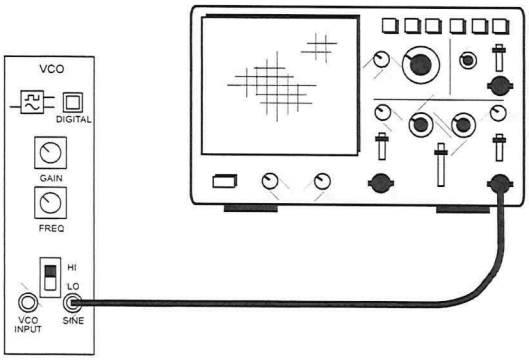


Figure 3

27. Use the PicoScope's measurement tools to help set the VCO module's output frequency to as close to 140kHz as you can obtain.

    Note 1: There's no need to adjust any of the scope's settings to observe a few cycles of this 4Vpp sinewave as a stable display as you might normally would.

    Note 2: If you can't set the VCO module's output frequency to 140kHz, set it to 130kHz instead but remember that you've done so when answering Question 9.

28. Modify the set-up as shown in Figure 4 below.

    Remember: Dotted lines show leads already in place.

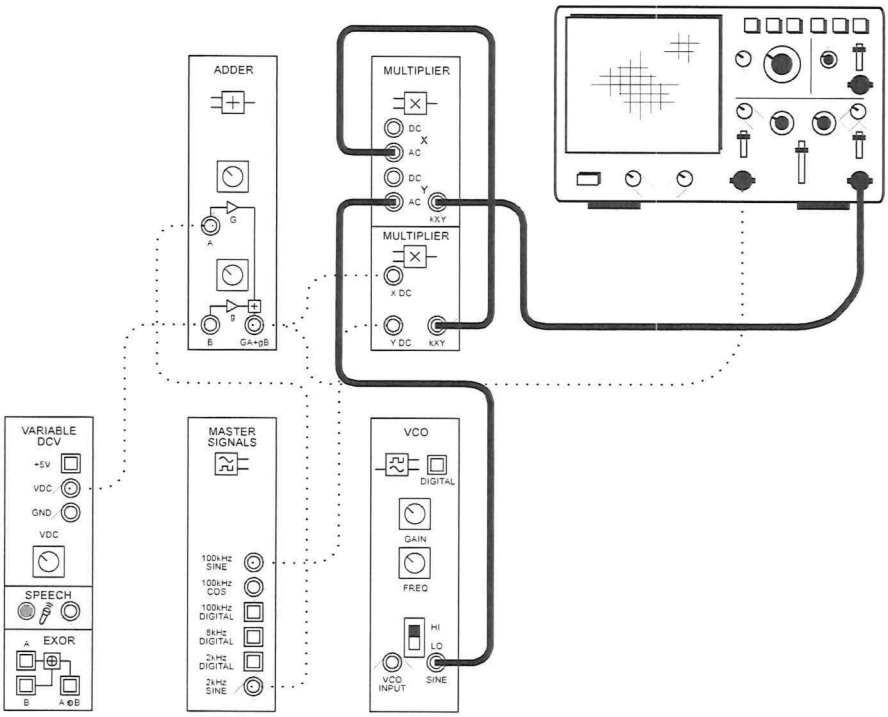


Figure 4

29. Set the scope's Channel B vertical scale to *±10V.*

30. Examine the signal on the Multiplier module's output.

    Note: You may find it helpful to pause the scope's acquisition to freeze the display.

Question 6
You're multiplying an AM signal with a 140kHz sinewave so you're not going to see an AM signal on Channel B. But, how would you describe this signal to another technician if you had to?

Question 7
According to the theory, for the given inputs to the Multiplier module, how many significant sinewaves does this signal consist of?

Question 8
According to the theory, are any of these sinewaves at 98kHz, 100kHz and 102kHz? Explain your answer.

Question 9
According to the theory, what are the frequencies of the significant sinewaves in the signal?

Ask the instructor to check your work before continuing.

31. Set the PicoScope to Spectrum Mode.

32. Examine the spectral composition of Multiplier module's output signal.

33. Identify the down-converted and up-converted versions of the original AM signal.

    Note: You may find it helpful to use the PicoScope's vertical cursors to identify them by looking for the two new carrier frequencies.

Question 10
The preliminary discussion mentions that the original sinewaves in a multiplication process are ostensibly lost. Why do you think the word "ostensibly" was used? **Tip:** The answer is in the spectrum of the Multiplier module's output.

_____

_____

_____

Ask the instructor to check your work before continuing.

Question 11
What do you think would happen to the carriers of these two AM signals if the VCO's frequency was increased (say, to 150kHz)? **Note:** See if you can figure this out for yourself without testing it.

- ☐ The carrier of both the down-converted and up-converted AM signals would decrease
- ☐ The carrier of both the down-converted and up-converted AM signals would increase
- ☐ The carrier of the down-converted AM signal would decrease while the carrier of the up-converted AM signal would increase
- ☐ The carrier of the down-converted AM signal would increase while the carrier of the up-converted AM signal would decrease

34. Test your answer to Question 11.

Ask the instructor to check your work before finishing.

analog
digital
biskit™

Emona-101 Trainer Lab Manual Volume 1 -
Experiments in Modern Analog and Digital Telecommunications.

Author: Barry Duncan

Emona Instruments Pty Ltd
86 Parramatta Road
Camperdown NSW 2050
AUSTRALIA

web: **www.ett101.com**
telephone: +61-2-9519-3933
fax: +61-2-9550-1378